

# Reconciling escape fractions and observed line emission in Lyman-continuum-leaking galaxies<sup>★</sup>

L. Ramambason<sup>1,2</sup>, D. Schaerer<sup>1,3</sup>, G. Stasińska<sup>4</sup>, Y. I. Izotov<sup>5</sup>, N. G. Guseva<sup>5</sup>, J. M. Vílchez<sup>6</sup>,  
R. Amorín<sup>7,8</sup>, and C. Morisset<sup>9</sup>

<sup>1</sup> Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, 1290 Versoix, Switzerland

<sup>2</sup> AIM, CEA, Université Paris-Saclay, Université Paris Diderot, Université de Paris, 91191 Gif-sur-Yvette, France  
e-mail: lise.ramambason@cea.fr

<sup>3</sup> CNRS, IRAP, 14 avenue E. Belin, 31400 Toulouse, France

<sup>4</sup> LUTH, Observatoire de Meudon, 92195 Meudon Cedex, France

<sup>5</sup> Bogolyubov Institute for Theoretical Physics, National Academy of Sciences of Ukraine, 14-b Metrolohichna Str., Kyiv 03143, Ukraine

<sup>6</sup> Instituto de Astrofísica de Andalucía, CSIC, Apartado de Correos 3004, 18080 Granada, Spain

<sup>7</sup> Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile

<sup>8</sup> Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile

<sup>9</sup> Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 106, 22800 Ensenada, BC, Mexico

Received 10 June 2020 / Accepted 17 September 2020

## ABSTRACT

**Context.** Finding and elucidating the properties of Lyman-continuum(LyC)-emitting galaxies is an important step in improving our understanding of cosmic reionization.

**Aims.** Although the  $z \sim 0.3\text{--}0.4$  LyC emitters found recently show strong optical emission lines, no consistent quantitative photoionization model taking into account the escape of ionizing photons and inhomogeneous interstellar medium (ISM) geometry of these galaxies has yet been constructed. Furthermore, it is unclear to what extent these emission lines can be used to distinguish LyC emitters.

**Methods.** To address these questions we construct one- and two-zone photoionization models accounting for the observed LyC escape, which we compare to the observed emission line measurements. The main diagnostics used include lines of [O III], [O II], and [O I] plus sulfur lines ([S II], [S III]) and a nitrogen line ([N II]), which probe regions of different ionization in the ISM.

**Results.** We find that single (one-zone) density-bounded photoionization models cannot reproduce the emission lines of the LyC leakers, as pointed out by earlier studies, because they systematically underpredict the lines of species of low ionization potential, such as [O I] and [S II]. Introducing a two-zone model, with differing ionization parameter and a variable covering fraction and where one of the zones is density-bounded, we show that the observed emission line ratios of the LyC emitters are well reproduced. Furthermore, our model yields LyC escape fractions, which are in fair agreement with the observations and independent measurements. The [O I]  $\lambda 6300$  excess, which is observed in some LyC leakers, can be naturally explained in this model, for example by emission from low-ionization and low-filling-factor gas. LyC emitters with a high escape fraction ( $f_{\text{esc}} \gtrsim 38\%$ ) are deficient both in [O I]  $\lambda 6300$  and in [S II]  $\lambda \lambda 6716, 6731$ . We also confirm that a [S II]  $\lambda \lambda 6716, 6731$  deficiency can be used to select LyC emitter candidates, as suggested earlier. Finally, we find indications for a possible dichotomy in terms of escape mechanisms for LyC photons between galaxies with relatively low ( $f_{\text{esc}} \lesssim 10\%$ ) and higher escape fractions.

**Conclusions.** We conclude that two-zone photoionization models are sufficient and required to explain the observed emission line properties of  $z \sim 0.3\text{--}0.4$  LyC emitters. This is in agreement with UV absorption line studies, which also show the co-existence of regions with high hydrogen column density (i.e., no escape of ionizing photons) and density-bounded or very low column density regions responsible for the observed escape of LyC radiation. These simple but consistent models provide a first step towards the use of optical emission lines and their ratios as quantitative diagnostics of LyC escape from galaxies.

**Key words.** galaxies: starburst – galaxies: high-redshift – dark ages, reionization, first stars – ultraviolet: galaxies

## 1. Introduction

Significant progress has been made in recent years in searches for and studies of star-forming galaxies, which emit copious amounts of Lyman continuum (LyC) radiation. This ionizing radiation can escape the interstellar medium (ISM) and ionize the surroundings, thus contributing to intergalactic ionizing radiation. If a sufficient population of such galaxies exists in the

early Universe, their presence could explain cosmic reionization, which is known to have started shortly after the Big Bang and to have ended  $\sim 1$  Gyr later, at  $z \sim 6$ .

After intense efforts to search for LyC emitters or LyC-leaking galaxies (abbreviated to “leakers”), several groups have identified such galaxies from low redshift up to  $z \sim 4$  (e.g., [Leitherer et al. 2016](#); [Izotov et al. 2016a,b, 2018a,b](#); [Vanzella et al. 2016, 2018, 2020](#); [de Barros et al. 2016](#); [Steidel et al. 2018](#); [Fletcher et al. 2019](#); [Bian et al. 2017](#); [Wang et al. 2019](#)) with LyC escape fractions,  $f_{\text{esc}}$ , ranging from a few percent up to  $\sim 72\%$  in some cases

<sup>★</sup> Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under 0102.B-0942(A).

(Izotov et al. 2018a; Vanzella et al. 2020). Although different methods and criteria have been used to search for these galaxies, some of the distinct observational features shared by many LyC emitters are strong Ly $\alpha$  emission with a double-peaked Ly $\alpha$  line profile (see Verhamme et al. 2015, 2017; Vanzella et al. 2020) and intense optical emission lines with a high [O III]/[O II]<sup>1</sup> ratio (Izotov et al. 2018b; de Barros et al. 2016; Vanzella et al. 2020). Following the suggestion by Jaskot & Oey (2013) and Nakajima & Ouchi (2014) that high values of [O III]/[O II] could pinpoint density-bounded galaxies, that is, LyC-emitting galaxies, this simple prediction was put to test by Izotov and collaborators, who selected compact star-forming galaxies with intense emission lines and [O III]/[O II] > 5, and who were thus able to find 11 LyC emitters at  $z \sim 0.3$ – $0.4$  with 100% success rate using the COS spectrograph on-board the *Hubble* Space Telescope (HST). The galaxies found in this way have properties comparable to the numerous galaxies at  $z > 6$ . They are strong producers of ionizing photons, show significant LyC escape ( $f_{\text{esc}} \sim 2$ – $72\%$ ), and are therefore good analogs of the sources of cosmic reionization (cf. Schaerer et al. 2016).

Despite the fact that high [O III]/[O II] was used as a basic selection criterion for these searches, it nevertheless turns out that  $f_{\text{esc}}$  does not well correlate with the [O III]/[O II] ratio (see Izotov et al. 2018b; Naidu et al. 2018; Bassett et al. 2019; Nakajima et al. 2020, and discussions therein). Similarly, Wang et al. (2019) recently selected some galaxies by their [S II] deficiency, which could indicate a density-bounded ISM, and were able to detect LyC emission from two low- $z$  galaxies with fairly low excitation ([O III]/[O II]  $\sim 1.3$ – $1.5$ ). However, these authors were not able to find a correlation between the [S II] deficiency and the LyC escape fraction of the known leakers. These findings are not entirely surprising, as it is well known that the [O III]/[O II] ratio depends also on several other factors, including in particular the ionization parameter and metallicity. Whether or not density-bounded models are applicable to known LyC emitters therefore remains to be examined quantitatively.

On the other hand, the detailed analysis of UV absorption lines (including the Lyman series lines of hydrogen and other low-ionization lines) of the  $z \sim 0.3$ – $0.4$  leakers of Izotov and collaborators undertaken by Gazagnes et al. (2018) and Chisholm et al. (2018) revealed the presence of high-column-density gas that is optically thick to LyC radiation and co-exists with lines of sight, which allows the escape of UV ionizing and nonionizing radiation from these galaxies. This implies that the ISM of these galaxies cannot be described by simple, one-zone photoionization models, and that a more complex ISM structure should be accounted for to consistently describe these objects. Therefore, testing one-zone photoionization models tailored to the properties (metallicities etc.) of the LyC emitters is required, as well as the exploration of more complex structures to consistently describe their emission line spectra. These are two of the main goals of our study.

Specifically, we would like to know whether or not simple – that is, single – 1D photoionization models can reproduce the observed emission-line properties of the LyC emitters, and if so, how well they do this. Comparing grids of photoionization models with observational diagrams of approximately 800 low-to moderate-metallicity star-forming galaxies, Stasińska et al. (2015) concluded that the strong lines indicate no strong LyC leakage (i.e.,  $f_{\text{esc}} \lesssim 10\%$ ) in galaxies with high [O III]/[O II],

because the observed emission line ratios are not compatible with density-bounded H II regions and therefore with a significant ionizing photon escape. The same authors also stressed the importance of considering high- and low-excitation lines. Furthermore, they conjectured that some leakage might occur because of a nebular covering factor smaller than unity, and that [OI] emission could come from optically thick clumps embedded in a density-bounded medium (Stasińska et al. 2015). Related to this, Plat et al. (2019) recently noted that low- $z$  LyC emitters show a higher [OI]/[O III] ratio than their ionization-bounded photoionization models, in contrast to expectations from density-bounded models where [OI]/[O III] should be reduced. To solve this discrepancy these latter authors invoke a contribution of hard radiation from active galactic nuclei (AGNs) or radiative shocks. However, no consistent model allowing for the observed leakage and escape fraction, including the geometrical constraints derived from the UV lines (as mentioned above), and explaining the intensities of the major emission lines has yet been made.

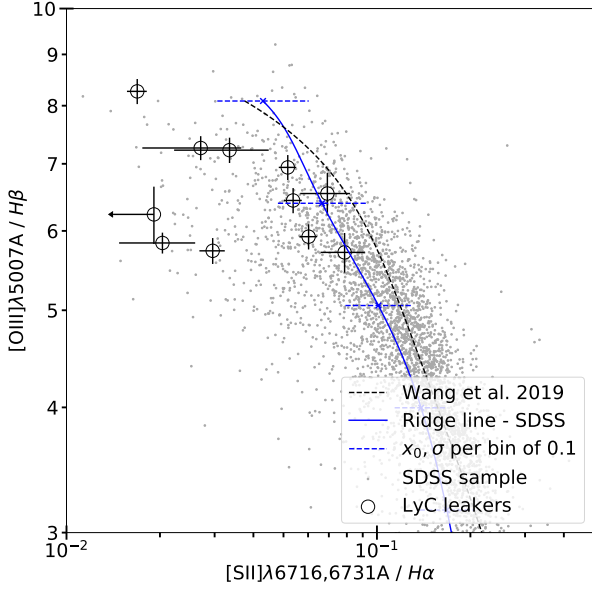
To progress further on these issues, here we compute both one- and two-zone photoionization models with metallicities tailored to those of the observed LyC emitters to quantitatively examine how such models compare to the various detailed observational constraints, including measured LyC escape fractions. First, we re-examine the most important observed emission line ratios of the  $z \sim 0.3$ – $0.4$  leakers, and compare them to a large sample of star-forming galaxies of comparable metallicity, serving as a meaningful reference sample, which is presumably dominated by nonleaking galaxies. We then compare the data to standard one-zone photoionization models to see the successes and failures of these models. Finally, because one-zone models cannot properly reproduce the data, we are led to construct two-zone models which show a lot of interesting features and provide useful insight into the ISM and emission-line properties of the LyC emitters.

If successful, consistent models should also allow us to better understand the origin of the different optical emission lines, their dependency on geometrical factors, leaking, and other factors. Hopefully, these models will lead to the development of a new diagnostic of LyC leakage using rest-frame optical emission lines that is applicable to other observations, especially of high- $z$  galaxies, for which many spectra will be obtained in the near future with the *James Webb* Space Telescope (JWST).

The need for multi-zone (or multi-component) models is not only of importance for LyC emitters. Indeed, numerous papers have pointed out the necessity to account for variations in density, ionization parameter, or geometry in order to simultaneously explain the observed strong emission line ratios of galaxies including different elements and ionization stages (see Stasińska & Leitherer 1996; Stasińska et al. 2006). Similarly, detailed studies of nearby galaxies with a large set of observations have clearly shown the limitations of one-zone models and the need for more complex and realistic representations (e.g., Stasińska & Schaerer 1999; Cormier et al. 2012, 2019; Lebouteiller et al. 2017). Here we explore some of these aspects with a focus on LyC-emitting galaxies.

This paper is organized as follows. In Sect. 2 we present our observational samples. Section 3 shows the main optical emission line diagrams used to compare to “classical” single-component photoionization models (both density- and ionization-bounded). We then construct two-component photoionization models including LyC leaking emission and compare them with observations (Sect. 4). Our results are discussed in Sect. 5. Section 6 summarizes and concludes the paper.

<sup>1</sup> In all the article we will use [O III] for [O III]  $\lambda 5007$ , [O II] for [O II]  $\lambda 3726, 3728$ , [OI] for [OI]  $\lambda 6300$ , [N II] for [N II]  $\lambda 6584$ , [S III] for [S III]  $\lambda 9068$  and [S II] for [S II]  $\lambda 6716, 6731$ .



**Fig. 1.** O3H $\beta$ –S2H $\alpha$  diagram of the LyC leakers (open circles) and comparison sample with metallicities  $12 + \log(\text{O}/\text{H}) = 7.6\text{--}8.2$  (gray dots). The plain blue line shows the ridge line of our SDSS sample, and the black dashed line is the ridge line from Wang et al. (2019) extrapolated to high [O III]/H $\beta$  values. Many leakers, especially the strongest ones, show a deficiency in [S II] with respect to this latter ridge line.

Throughout this paper, we assume a solar chemical composition following Asplund et al. (2009) where  $\log(Z/Z_{\odot}) = 12 + \log(\text{O}/\text{H}) - 8.69$ .

## 2. Optical emission line properties of LyC emitters

In the present work we study eleven  $z \sim 0.3\text{--}0.4$  Lyman continuum emitters discovered with the COS spectrograph on-board HST and whose properties have been amply discussed in several papers (Izotov et al. 2016a,b, 2018a,b; Schaerer et al. 2016, 2018; Verhamme et al. 2017; Gazagnes et al. 2018; Chisholm et al. 2018). The metallicities of these galaxies are in the range  $12 + \log(\text{O}/\text{H}) = 7.65\text{--}8.16$ , with a median oxygen abundance of  $12 + \log(\text{O}/\text{H})$  of 7.91.

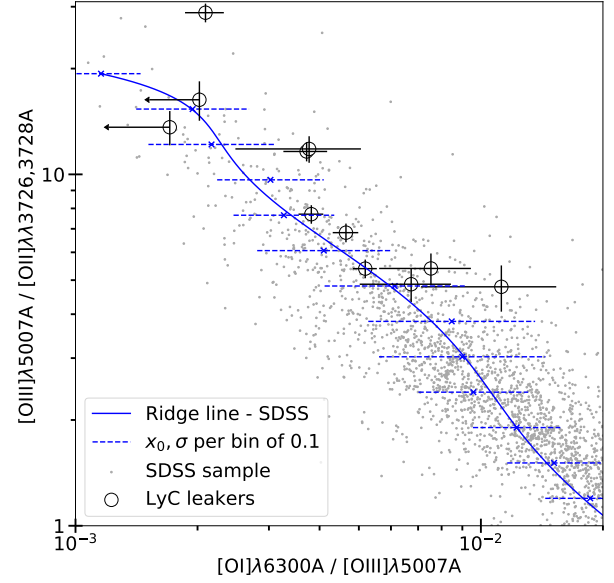
These sources represent currently the best sample of local LyC emitters, for which a wide range of observations is available, including LyC, UV, and optical spectroscopy. Furthermore, Gazagnes et al. (2018, 2020) provided a detailed analysis of their UV absorption lines (including Lyman series lines and other low-ionization absorption lines), which provides important geometrical constraints on the ISM of these galaxies. Here we analyze the optical emission line properties of these LyC emitters.

Before discussing the observed emission line properties of the LyC emitters, we briefly summarize the observational data used, including the main galaxies of interest (11 LyC emitters) and an appropriate comparison sample.

### 2.1. Observational data

#### 2.1.1. Lyman continuum emitters

The emission line measurements from the SDSS spectra of the LyC emitters are reported in the discovery papers (Izotov et al. 2016a,b, 2018a,b). We recently obtained new VLT spectra of 5 of the 11 leakers (J0925+1403, J0901+2119, J1011+1947, J1154+2443 and J1442–0209) with the X-shooter spectrograph.



**Fig. 2.** O32–O13 diagram.

We use the improved emission line measurements from these spectra, and in particular also the [S III]  $\lambda 9068$  measurements, which are present in the near-IR spectra of these galaxies. A more detailed description of these observations and results are presented in Guseva et al. (2020). All emission line ratios reported here refer to extinction-corrected ratios.

#### 2.1.2. Comparison sample

For comparison we use a sample of compact star-forming galaxies from the Sloan Digital Sky Survey (SDSS) Data Release 14 (Abolfathi et al. 2018) analyzed in earlier publications (Guseva et al. 2019), and for which the [O III]  $\lambda 4363$  line is detected with an accuracy better than  $4\sigma$ , allowing direct abundance determinations using the  $T_e$  method. In practice, from the list of 5607 galaxies with direct metallicity measurements, we retain 4571 galaxies in the metallicity range  $12 + \log(\text{O}/\text{H}) = 7.6\text{--}8.2$ , for comparison with the LyC emitters discussed here.

### 2.2. Optical emission line properties of LyC emitters

The emission line properties of the LyC emitters and the comparison sample are shown in Figs. 1, 2, and 4. By selection, all the galaxies studied here, including the LyC emitters, have emission lines ratios compatible with those of star-forming galaxies, that is, they fall in the range of stellar photoionization sources in the classical diagrams by Veilleux & Osterbrock (1987) used for spectral classification (including the [N II] BPT diagram introduced by Baldwin et al. 1981). Also, by selection, the sources studied here have a relatively low metallicity, which implies a fairly high excitation, reflected for example by a high [O III]/H $\beta$  ratio. However, at a finer level, the LyC emitters show some interesting distinctive properties, which we now discuss.

#### 2.2.1. A [S II] deficiency in LyC emitters

In Fig. 1 we show a classical diagnostic (O3H $\beta$ –S2H $\alpha$ ) involving the [S II]/H $\alpha$  ratio, which was used by Wang et al. (2019) to select LyC-emitter candidates among more metal-rich and therefore lower excitation (lower [O III]/H $\beta$ ) sources. Clearly several,

**Table 1.** Coefficients of the ridge lines in classical diagrams (Figs. 1, 4 and 8), determined for the entire galaxy sample and described by Eq. (1).

$x$	[N II]/H $\alpha$	[S II]/H $\alpha$	[O I]/H $\alpha$
$\alpha_0$	-2.40	-2.00	-3.27
$\alpha_1$	5.22e-1	3.78e-1	4.33e-1
$\alpha_2$	-5.74e-2	-3.98e-2	-4.33e-2
$\alpha_3$	3.14e-3	2.11e-3	2.24e-3
$\alpha_4$	-9.77e-4	-6.40e-3	-6.68e-3
$\alpha_5$	1.80e-4	1.16e-4	1.20e-4
$\alpha_6$	-1.94e-4	-1.23e-4	-1.28e-4
$\alpha_7$	1.14e-4	7.16e-3	7.48e-3
$\alpha_8$	-2.79e-3	-1.74e-3	-1.83e-3

**Notes.** Here  $y = [\text{O III}] \lambda 5007/\text{H}\beta$ .

**Table 2.** Coefficients of the ridge lines in oxygen and sulfur diagrams (Figs. 2, 5 and 9), described by Eq. (1).

$x$	[O I]/[O III]	[S II]/[S III]
$\alpha_0$	-1.67	4.13e-1
$\alpha_1$	-1.07	-5.15e-1
$\alpha_2$	4.68e-1	-1.54
$\alpha_3$	1.46	1.16e-1
$\alpha_4$	-9.67e-1	1.30
$\alpha_5$	-5.33	-9.20e-1
$\alpha_6$	4.36	1.75e-2
$\alpha_7$	2.01	-1.28e-2
$\alpha_8$	-1.84	3.31e-1

**Notes.** Here  $y = [\text{O III}]/[\text{O II}]$ .

if not a majority, of the eleven confirmed leakers show fairly weak [S II]  $\lambda\lambda 6716, 6731$  emission, or [S II]/H $\alpha$  ratios close to or below the typical value for galaxies from our comparison sample at the same [O III]  $\lambda 5007/\text{H}\beta$  ratio. To quantify this comparison we derive an SDSS “ridge line” which marks the median location of normal star-forming galaxies in the same range of O/H (metallicity) as our LyC emitters. This was obtained by binning the data in logarithmic bins of 0.1 dex along the [O III]/H $\beta$  axis with each bin having at least ten galaxies. Each bin was then fitted by a Gaussian distribution of a center  $x_0$  and a standard deviation  $\sigma$ . The ridge line defined by the  $x_0$  was then fitted with an eight-order polynomial,

$$\log x = \sum_{n=0}^8 \alpha_n (\log y)^n, \quad (1)$$

where  $y = [\text{O III}] \lambda 5007/\text{H}\beta$ , and the coefficients  $\alpha_n$  are reported in Table 1. Table 2 presents the coefficients of the ridge lines derived in other diagnostic plots  $x$  vs.  $y = [\text{O III}]/[\text{O II}]$ .

For subsequent discussions, we also calculate  $\Delta[\text{S II}]$ , the lateral shift from the ridge line,

$$\Delta[\text{SII}] = \log ([\text{SII}]/\text{H}\alpha) - f(\log ([\text{OIII}]/\text{H}\beta)), \quad (2)$$

to quantify deviations of individual galaxies from the average observed line ratio of the entire comparison sample. With this definition  $\Delta[\text{S II}] < 0$  indicates a deficit in [S II]/H $\alpha$  with respect to the average at the same [O III]  $\lambda 5007/\text{H}\beta$  ratio. We note that the sulfur abundance ratio (S/O) is constant and shows a fairly small scatter over a wide range of metallicities, even beyond the

range of O/H considered here (see, e.g., Guseva et al. 2020). Therefore, the mean trends described by this ridge line should be quite independent of the exact S/O abundance.

As shown in Fig. 1, five of our LyC emitters are within  $1\sigma$  of the ridge line, whereas the rest are systematically offset to lower [S II]/H $\alpha$  ratios, that is, they show an [S II] deficit, an empirical finding already pointed out by Wang et al. (2019). Several reasons could be invoked to explain such a behavior: S/H abundance variations; a decrease of [S II] due to the loss of the outer, neutral regions related to the escape of LyC radiation; a lower contribution of diffuse ionized gas (DIG), which is in general known to “boost” [S II] emission (cf. Sanders et al. 2020); or possibly other explanations. The [S II] deficit and its link with the observed LyC escape fraction are discussed below, providing a more quantitative description of the relationship (Sect. 4.4).

## 2.2.2. An excess of [O I]/[O III]

Another peculiarity of the LyC emitters is found in the diagram comparing the three ionization stages of oxygen traced by the optical line ratios [O III]/[O II] and [O I]/[O III], as shown in Fig. 2 (denoted the O32–O13 diagram hereafter). Indeed, naively one would expect to find a deficit of [O I] emission in LyC emitters as they should be highly ionized and possibly lacking neutral regions where [O I] is emitted. This is also shown by simple density-bounded photoionization models (see Figs. 5 and 6, and Stasińska et al. 2015). However, as also noted by Plat et al. (2019), [O I] is detected in 9 out of the 11 leakers of this study, and the observed [O I]/[O III] ratios of the LyC emitters are surprisingly high, shifted to even higher than average values when compared to the ridge line defined by the comparison sample in Fig. 2. This empirical finding clearly shows the presence of neutral oxygen in the LyC leakers, which already indicates that the ISM in these galaxies can probably not be described by simple density-bounded models, and whose presence and intensity needs to be explained quantitatively by models.

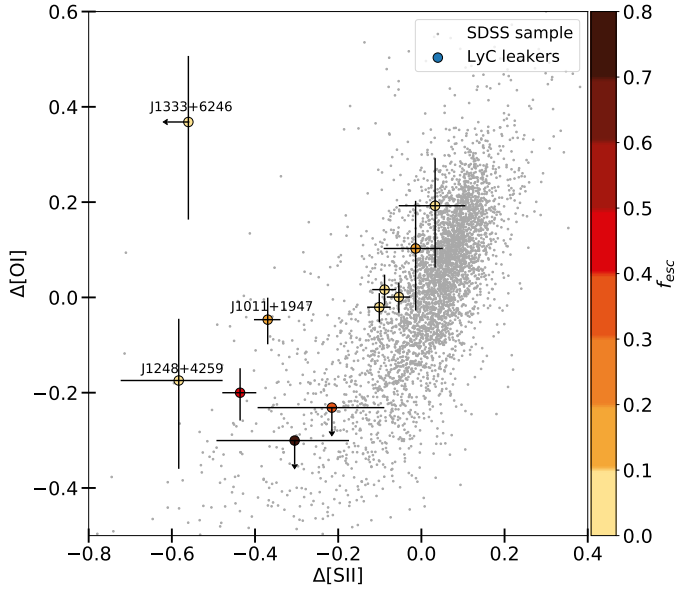
## 2.2.3. A shift in low-ionization line properties

In Fig. 3, we show the displacement relative to the ridge lines in O3H $\beta$ –O1H $\alpha$  and O3H $\beta$ –S2H $\alpha$  diagrams (see Cols. 2 and 3 in Table 1). In this plane, we see a clear shift from the SDSS comparison sample of presumably nonleaking galaxies. Interestingly, the three most leaking galaxies (the three reddest circles,  $f_{\text{esc}} \geq 38\%$ ) in our sample stand in the lower-left corner with low  $\Delta[\text{S II}]$  and  $\Delta[\text{O I}]$  ( $< -0.2$  dex). On the other hand, most of the small leakers in our sample (yellow circles with  $f_{\text{esc}} \lesssim 10\%$ ) exhibit a moderate deficiency in [S II] ( $-0.2 < \Delta[\text{S II}] < 0$ ) and [O I] excesses ( $\Delta[\text{O I}] \geq 0$ ). However, we note that three galaxies in our sample with a large deficit in [S II] ( $\Delta[\text{S II}] \lesssim -0.3$ ) have intermediate (J1011+1947,  $f_{\text{esc}} = 11.4\%$ ) to low escape fractions (J1248+4259, J1333+6246, labeled on the plot). One of them (J1333+6246) strikingly stands apart with a nondetection of the [S II] line but the largest [O I] excess in the sample. A more detailed analysis of these shifts in [O I] and [S II], including a comparison to models, is given in Sect. 4.4.

## 2.2.4. Other emission-line peculiarities of LyC emitters

In the [N II] BPT diagram (O3H $\beta$ –N2H $\alpha$ , Figs. 4 and 8) it is possible to note a slight offset of most of the leakers with respect to the average of the comparison sample. Likely related to this, Guseva et al. (2020) noted that the N/O abundance of LyC emitters is again higher than the average of star-forming galaxies at





**Fig. 3.**  $\Delta[\text{O I}]$  vs.  $\Delta[\text{S II}]$  derived from classical diagrams  $\text{O3H}\beta\text{--O1H}\alpha$  and  $\text{O3H}\beta\text{--S2H}\alpha$ . The three strongest leakers (J1256+4509, J1154+2443 and J1243+4646, by increasing order of  $f_{\text{esc}}$ ) are represented by orange to brown circles in the lower part of the plot. The other galaxies discussed in this paper are labeled. The LyC emitters show a clear displacement in this plane compared to the comparison sample.

the same O/H abundance. The origin of this behavior is currently unknown.

In another diagram discussed further below, relating  $[\text{O III}]/[\text{O II}]$  to  $[\text{S II}]$   $\lambda\lambda 6716, 6731/[\text{S III}]$   $\lambda 9068$  (Fig. 5), we note that four out of five LyC emitters, for which we have near-IR measurements of  $[\text{S III}]$   $\lambda 9068$ , appear also shifted from the average  $[\text{S II}]$   $\lambda\lambda 6716, 6731/[\text{S III}]$   $\lambda 9068$  ratio in a similar fashion as for  $[\text{O I}]/[\text{O III}]$   $\lambda\lambda 4959, 5007$ . More measurements will show if this behavior is systematic for LyC leakers or not.

Finally, Izotov et al. (2017) proposed that LyC-emitting galaxies show peculiar line ratios of weak He I lines. As shown by Guseva et al. (2020), our X-shooter observations confirm the trend suggested by this study.

To gain further insight into the observed behavior of the emission line ratios we now proceed to comparisons with photoionization models, first with simple and then more complex geometries.

### 3. Comparison to classical photoionization models

The first natural step in a comparison between the observed emission line properties of LyC emitters is to examine how simple 1D photoionization models allowing for varying degrees of ionizing photon escape behave and compare with the data.

#### 3.1. The models: BOND models from the 3MdB database

To examine the power of simple 1D photoionization models, we used Cloudy photoionization models assembled in the BOND (Bayesian Oxygen and Nitrogen abundance Determinations) grid of models (Vale Asari et al. 2016), which is available on the Mexican Million Models database (3MdB, Morisset et al. 2015). This grid was designed to be used with the BOND code, a Bayesian code, which simultaneously derives oxygen and nitrogen abundances in giant H II regions. One of the specificities of the BOND code is that it is independent of any assumption on

relations between O/H and N/O. Therefore, the grid spans a wide range in O/H, N/O, and ionization parameter  $U$ , and covers different starburst ages and nebular geometries (spherical bubble or thin shell). The spectral energy distribution (SED) of the ionizing radiation is obtained from the population synthesis code PopStar (Mollá et al. 2009) for a Chabrier (Chabrier 2003) stellar initial mass function (IMF) between 0.15 and  $100 M_{\odot}$ . Varying starburst ages describe variations in the ionizing radiation field hardness, which arise due to the evolution of the stellar populations ionizing the H II regions. A subset of the BOND model grid has been used by Stasińska et al. (2015), who also studied the effect of LyC leakage. In the current study, we use models from the 3MdB\_17 database<sup>2</sup> (ref = “BOND\_2”), which is an updated version of 3MdB using Cloudy v17.02 (Ferland et al. 2017) and covering a somewhat refined grid.

For our study, we used a sub-grid of BOND models, with parameters appropriate for the LyC-emitting galaxies and the comparison sample. Specifically, the metallicity range is limited to  $12 + \log(\text{O}/\text{H}) = 7.6\text{--}8.2$  (with a step of 0.2), covering the observed metallicities of the LyC emitters. The ionization parameter  $\log U$  varies between  $-4$  and  $-1$  (with a step of 0.25), the age between 1 and 3 Myr (in steps of 1 Myr). The models are spherical filled bubbles with a fixed density of  $100 \text{ cm}^{-3}$ , and abundance ratios of  $\text{N}/\text{O} = 10^{-1}$  and  $\text{S}/\text{O} = 10^{-1.66}$ . The choice of a spherical filled bubble geometry influences the ionization structure of the modeled H II regions: based on Stasińska et al. (2015), we expect thin shell models to produce fainter emission from low-ionization-potential species (e.g.,  $[\text{O I}]$ ,  $[\text{S II}]$ ) with respect to higher ionization potential species (e.g.,  $[\text{O III}]$ ,  $[\text{S III}]$ ). Specifically, this would induce a lateral shift to the left-hand side in the  $\text{O32--O13}$  and  $\text{O32--S23}$  diagnostics plots (Figs. 5 and 9), which we examine in Sect. 3.2.

To describe density-bounded models, we used a parameter  $H_{\beta} \text{frac}$  which cuts the nebula at different radii corresponding to a range of 10–100% of the total  $H_{\beta}$  luminosity<sup>3</sup>. This parameter corresponds to the fraction of  $H_{\beta}$  luminosity emitted by the entire modeled galaxy relative to the  $H_{\beta}$  luminosity that would be emitted by the same model if it was ionization-bounded. Since the  $H_{\beta}$  emissivity is roughly constant throughout the H II region, this quantifies the ratio of the radius of the truncated, density-bounded model to the Strömgren radius of its ionization-bounded analog. We directly translate  $H_{\beta} \text{frac}$  into the corresponding LyC escape fraction by defining  $f_{\text{esc}}(H_{\beta}) = 1 - H_{\beta} \text{frac}$  which we use as a proxy to estimate how much leakage is allowed by a given density-bounded model. We refer to the models used here as simple, one-zone models.

#### 3.2. The spectral diagnostics

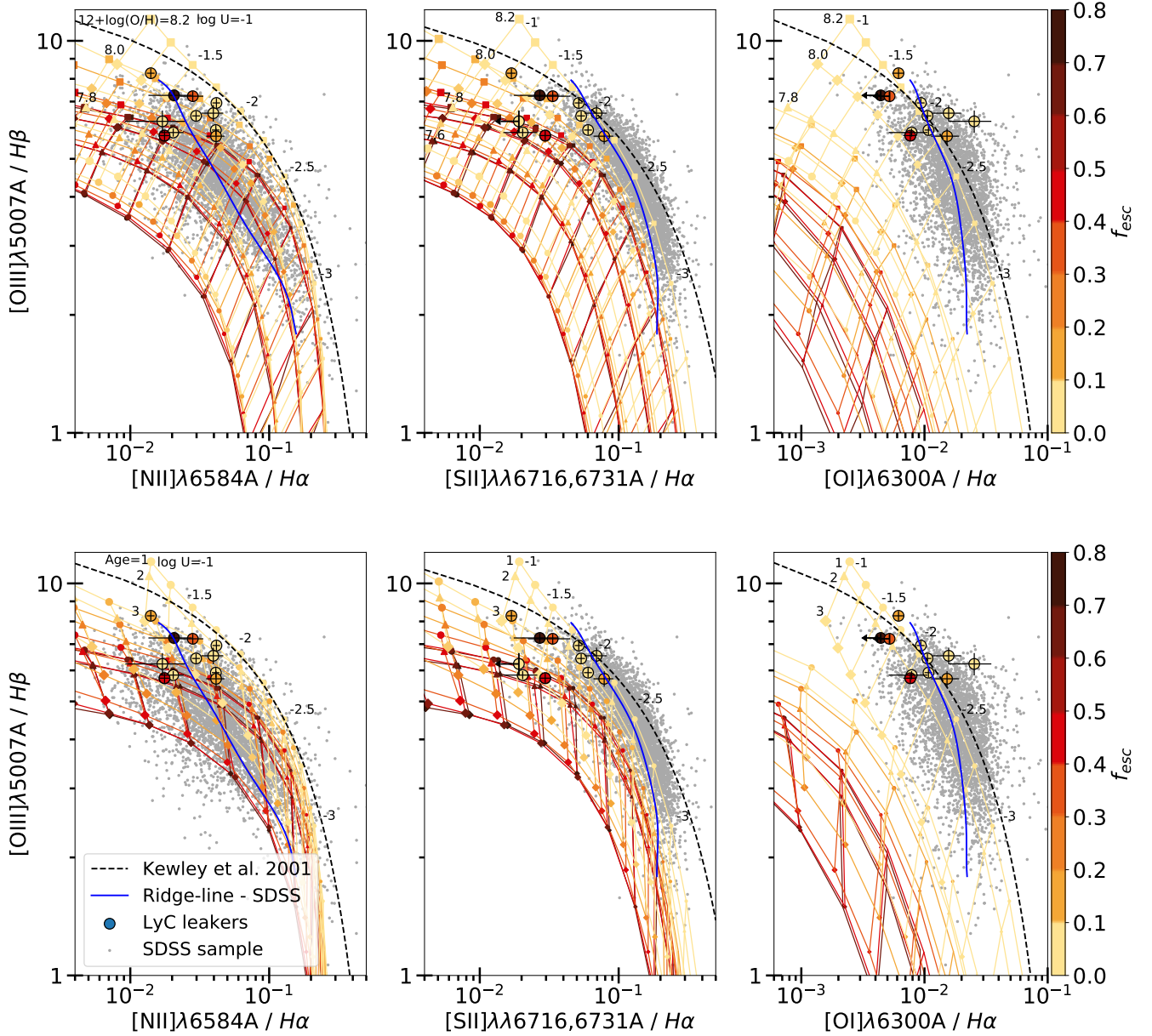
We now compare the one-zone photoionization models to the observations using the classical diagrams introduced by Veilleux & Osterbrock (1987) for spectral classification, the above-mentioned  $\text{O32--O13}$ , and other diagrams.

##### 3.2.1. Classical diagrams

Figure 4 shows the LyC emitters, the comparison sample, and the one-zone models in the classical diagrams from Veilleux & Osterbrock (1987), including the  $[\text{N II}]$  BPT diagram. As before,

<sup>2</sup> The 3MdB website describes the 3MdB and 3MdB\_17 databases at <https://sites.google.com/site/mexicanmillionmodels/>

<sup>3</sup> Since the study of Vale Asari et al. (2016), the grid in  $H_{\beta} \text{frac}$  has been refined and is now available for steps of 10%.



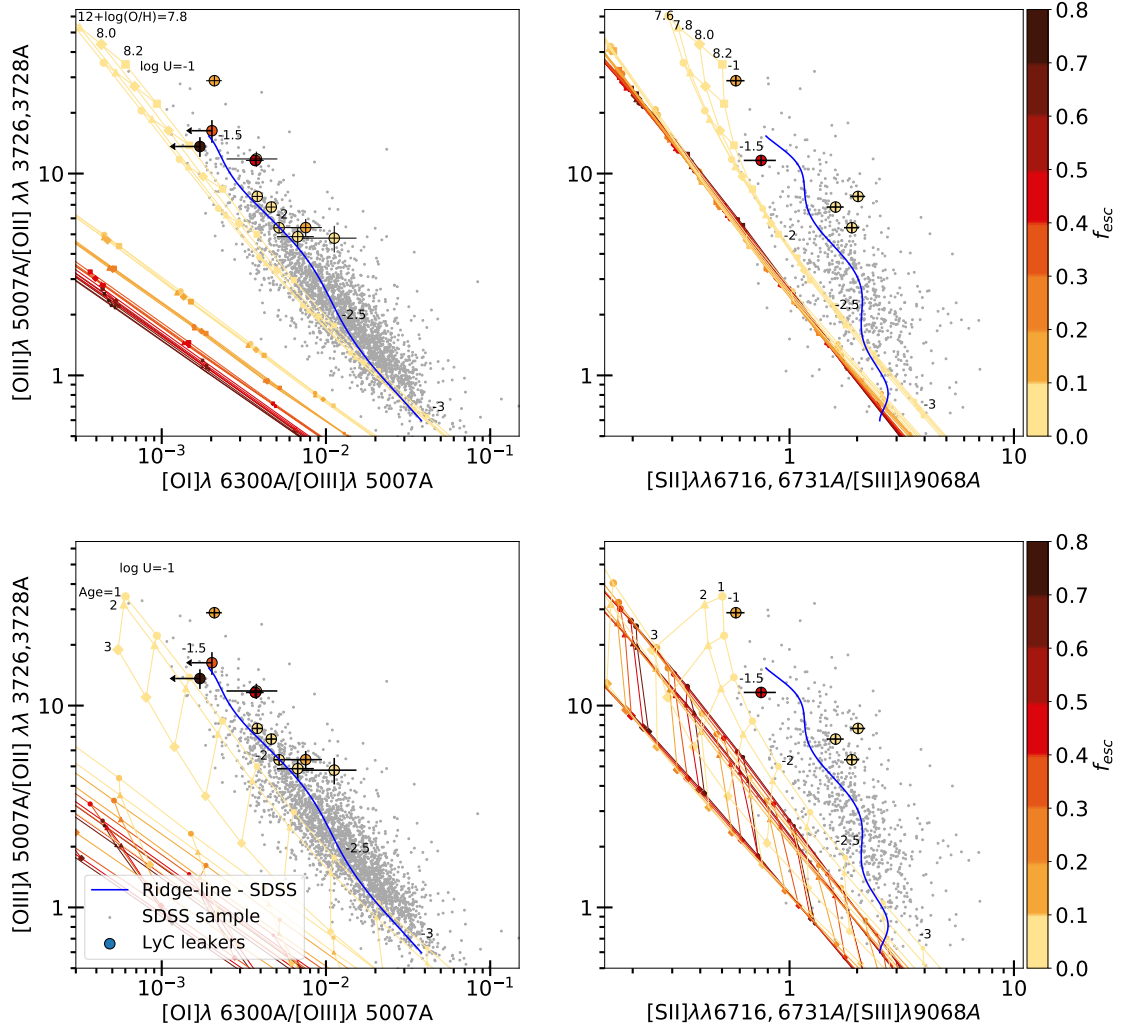
**Fig. 4.** Classical diagrams with BOND models and observations. Photoionization models are represented by the colored symbols: the colorbar represents the variation of the escape fraction and the size of symbol is proportional to  $\log(U)$ . The shape either represents the variation of metallicity (*first row*: for the age of 1 Myr) or stellar age (*second row*: for the metallicity of 8.2). The observed data points (gray dots) show the SDSS galaxy sample in the restricted metallicity range,  $12 + \log(\text{O}/\text{H}) = 7.6\text{--}8.2$ . The plain blue lines show the ridge lines of the entire sample. Dashed lines correspond to the AGN delimitations from Kewley et al. (2001).

we also indicate the ridge line, describing the location of the average SDSS galaxies in the corresponding plots (cf. Sect. 2.2). For illustration, dashed lines indicate the theoretical delimitations between star-forming galaxies (below) and AGN (above), according to the models of Kewley et al. (2001).

Overall, the ionization-bounded ( $f_{\text{esc}} = 0$ ) photoionization models cover reasonably well the observed line ratios of the entire galaxy sample including the LyC emitters. The  $\text{O3H}\beta\text{--N2H}\alpha$  diagram is better covered by these models than the  $\text{O3H}\beta\text{--S2H}\alpha$  and  $\text{O3H}\beta\text{--O1H}\alpha$  diagrams. For example, the ridge line in the  $[\text{S II}]\lambda\lambda 6716, 6731\text{A} / \text{H}\alpha$  diagram is located outside the parameter space examined here; discussing the origin of these offsets is beyond the present scope. The position of the LyC leakers indicates high ionization parameters and young stellar ages of the ionizing populations; the latter is in good agreement with the young ages found from the observed

UV spectral features and from spectral modeling of these galaxies (see, e.g., Izotov et al. 2019).

The second most important conclusion from Fig. 4 is that the domain of the emission line ratios observed in the LyC emitters is only covered by ionization-bounded photoionization models, that is, models with  $f_{\text{esc}} = 0$  or clearly only for low values of  $f_{\text{esc}} < 0.1$ , which is in contradiction with the observed escape fractions reaching up to  $\sim 70\%$  for some of the leakers. As already shown by Stasińska et al. (2015), increasing the escape fraction in these simple, one-zone 1D photoionization models strongly decreases the  $[\text{S II}]$  and  $[\text{OI}]$  emission, which results from trimming the outer parts of the  $\text{H II}$  region. This shows that overall the observed LyC emitters (or at least those with a significant LyC escape fraction) cannot be reproduced by simple density-bounded models, as these systematically underpredict



**Fig. 5.** O32–O13 and O32–S23 diagnostics with BOND models. Photoionization models are represented by the colored symbols: the colorbar represents the variation of the escape fraction and the size of symbols is proportional to  $\log(U)$ . The shapes of model dependencies reflect the variations either with metallicity or stellar age and are shown and labeled in the *top* and *bottom* panels, respectively.

lines of species with low ionization potentials, such as [O I] and [S II].

### 3.2.2. Diagnostics of the low-ionization regions

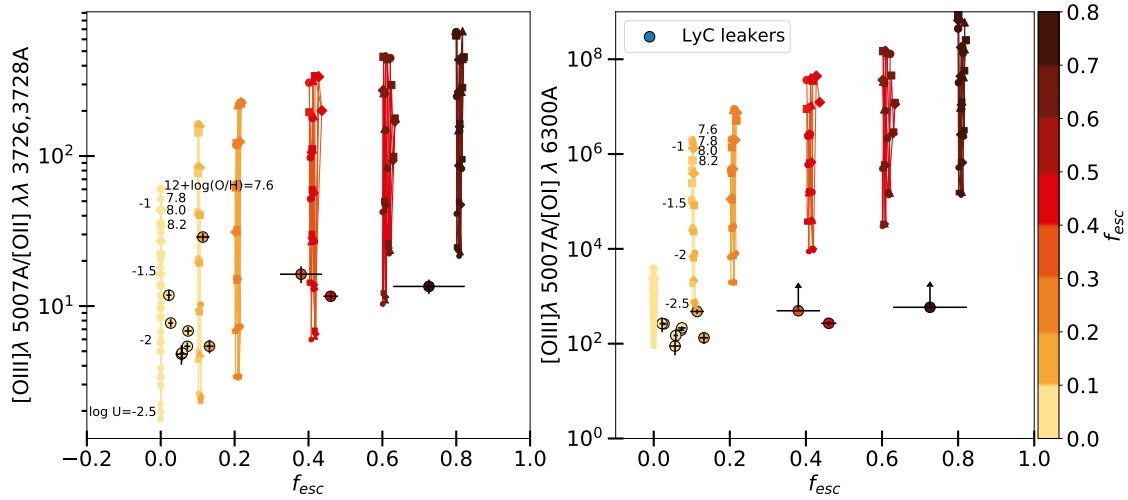
In Fig. 5 we examine complementary and probably more powerful emission line diagnostics to probe the ionization structure and therefore the different zones of the ionized ISM in the LyC emitters. By focusing on different ionization stages of the same element, we also avoid dependencies on elemental abundance ratios and metallicity, at least to first order. These O32–O13 and O32–S23 diagnostics highlight two main points that motivate the introduction of more complex models in the Sect. 4:

(1) Although the oxygen line ratios (O32–O13) predicted by simple ionization-bounded models reproduce the SDSS sequence with ionizing spectra of young populations ( $\sim 1$ – $2$  Myr) quite well, the models are not able to account for galaxies with a certain [O I]  $\lambda 6300$  excess, that is, a higher-than-average [O I]/[O III] ratio at a given [O III]/[O II] ratio, which is observed for the LyC emitters and other sources (see Fig. 5 left). For [S II]/[S III], shown on the right panel of this figure, the deficiency of the models is even more striking, as they fail to reproduce the bulk of the observations. The two discrepancies might

arise from the fact that the models do not take into account a DIG component, which can significantly contribute to the emission of [O I] and [S II] (Sanders et al. 2017, 2020). Alternatively, these discrepancies could also be attributed to the presence of shocks, which are not included in our models (see, e.g., Plat et al. 2019) because of problems with atomic data for sulfur (Izotov et al. 2006; Kewley & Dopita 2002; Kewley et al. 2019), or others. The sulfur line ratio [S II]/[S III] is also sensitive to more subtle effects including stellar winds, outflows, and photoevaporation of photo-dissociation regions (PDRs) in Galactic dense gaseous environments (e.g., Westmoquette et al. 2013; McLeod et al. 2015).

(2) The density-bounded models allowing for LyC photon escape do not at all reproduce the observed oxygen and sulfur line ratios of the leakers, as already noted above. On the contrary, while leakers populate the upper part of the SDSS sequence, the density-bounded models produce very low emission in [O I] and [S II], which drives them in the opposite direction to the bottom-left corner of our diagnostic plots. Figure 6 represents the [O III]/[O II] and [O III]/[O I] ratios predicted by the models as a function of their escape fraction. This diagram emphasizes the discrepancy between [O III]/[O I] and  $f_{\text{esc}}$ , showing that [O III]/[O I] is overestimated by several orders of magnitude.





**Fig. 6.** Observed and predicted  $[\text{O III}]/[\text{O II}]$  (left) and  $[\text{O III}]/[\text{O I}]$  (right) line ratios as a function of the LyC escape fraction, comparing the observed LyC emitters and the one-zone photoionization models with  $\log(U) \gtrsim -2.5$ . We note the very high range of the predicted  $[\text{O III}]/[\text{O I}]$  ratio.

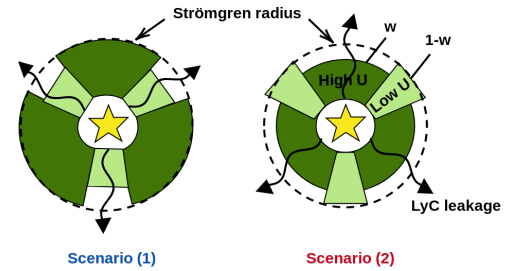
This clearly shows that moderate to high escape fractions cannot be produced by simple density-bounded one-zone models without significantly decreasing the weak line emission of  $[\text{O I}]$  and  $[\text{S II}]$ .

In short, although density-bounded models would explain the Lyman-continuum-escaping photons, they clearly do not match the spectral signatures arising from the outer part of the H II region and observed in LyC emitters. On the other hand, while ionization-bounded models can describe the overall properties of most galaxies (the comparison sample) fairly well – except for sulfur line ratios –, they underpredict the emission from low-ionization species (e.g., the  $[\text{O I}]$  line) in the LyC emitters, and most importantly, they do not account for the escape of LyC photons. These failures of simple, one-zone photoionization models clearly indicate that more complex models are required to describe the observed emission line properties of the LyC emitters. We therefore examined composite photoionization models based on the combination of two distinct zones with different properties.

## 4. Composite “two-zone” photoionization models

### 4.1. Motivation and method

As demonstrated above (Sect. 3.2), single-component (one-zone) photoionization models fail to simultaneously reproduce the observed line ratios and escape fractions of LyC leakers. To answer our problem (1) from Sect. 3.2.2 (insufficient emission in low-ionization lines), we introduce a component of low ionization (low  $U$ ) that would mimic a DIG contribution in star-forming galaxies. The relative contribution of this component is described by a factor  $\omega$  that varies from one galaxy to another. In the case of LyC emitters, although they are expected to be relatively poor in DIG (see Oey et al. 2007), this component could effectively represent the effect of low-ionization channels carved into the ISM through which some LyC photons escape, and which may contain gas with a low-volume filling factor, such as clumps or filaments. As an answer to problem (2) from Sect. 3.2.2 (inconsistency between the escape fraction and the presence of low-ionization lines), we include in our two-zone model the possibility for LyC photon escape from one region, while the second remains ionization-bounded.



**Fig. 7.** Schematic illustration of our two-zone photoionization models for LyC emitters, combining an ionization region with density-bounded regions chosen to have different ionization parameters (“high” and “low”  $U$ ). In the “density-bounded channels” scenario, ionizing photons escape from the low- $U$  region, while in the “density-bounded galaxy” scenario, they escape from the high- $U$  region with  $f_{\text{esc}} = 0$  for the nonleaking region and  $f_{\text{esc}}$  between 0% and 80% for the leaking region.

A two-zone, “picket-fence”-type model, was also suggested based on other observations. Gazagnes et al. (2018, 2020) and Chisholm et al. (2018) analyzed the UV absorption lines of hydrogen and low-ionization metal lines of the LyC emitters studied here. Their work clearly shows the presence of absorption lines formed in the neutral gas, with column densities that are too high for the escape of LyC photons. Furthermore, the depths of the saturated absorption lines show that channels of low column density must be present in these galaxies, allowing LyC escape. For our purposes, their model is basically equivalent to our two-zone model. As we show below, such an idealized two-zone model can also capture and describe the observed optical emission lines of the LyC emitters.

In practice, we have considered two different scenarios that are represented in Fig. 7. Each scenario is based on a linear combination of two models from the BOND database (see Sect. 3.1), one with a high-ionization parameter weighted by a factor  $\omega$ , and one with a low  $U$  weighted by a factor  $(1-\omega)$ . In each scenario, we allow only one component to be leaking.

The high- $U$  component corresponds to the main body of the galaxy which is highly ionized and homogeneous (filling factor  $\epsilon \sim 1$ ). We briefly discuss here the motivation for introducing a low-ionization component. The BOND models are



parameterized by an input volume-averaged ionization parameter that, for a given geometry, is proportional to the ionization parameter of the corresponding ionization-bounded model at the Strömgren radius (see Eq. (4) in Stasińska et al. 2015). For ionization-bounded models, the ionization parameter at the Strömgren radius is defined by the following equations:

$$U(r = R_S) = \frac{Q}{4\pi R_S^2 n c}, \text{ and } R_S = \left( \frac{3Q}{4\pi \alpha(T_e) n^2 \epsilon} \right)^{1/3}, \quad (3)$$

where  $Q$  is the number of ionizing photons,  $n$  the density of hydrogen,  $\alpha(T_e)$  the recombination coefficient of hydrogen, and  $\epsilon$  the filling-factor for an inhomogeneous medium ( $0 \lesssim \epsilon \lesssim 1$ ). Equation (3) leads to:

$$U(r = R_S) = A(Qn\epsilon^2)^{1/3}, \quad (4)$$

where  $A$  is almost constant ( $A \approx 1.3 \times 10^{-21} \text{ cm s}^{1/3}$  for  $T_e = 10^4 \text{ K}$ ), which means that both low-density and clumpy mediums ( $\epsilon \ll 1$ ) will correspond to a model with a low input ionization parameter. We note that the density-bounded models created by imposing a cut in radius to the full ionization-bounded model do not necessarily have a mean ionization parameter that matches their input ionization parameter. The two components do not have to share exactly the same Strömgren radius but one can ensure that the physical sizes are compatible by adjusting the filling factor of one component to match the size of the other. We now further describe the scenarios in Fig. 7.

**Density-bounded channels (DBC).** Ionization-bounded galaxy with density-bounded channels. The galaxy is ionization-bounded but LyC photons can escape through channels, which are matter-bounded. Such holes could be carved in the ISM by supernovae explosions or could result from hydrodynamical fragmentation effects induced by turbulence or stellar feedback (e.g., Kakiichi & Gronke 2019; Hogarth et al. 2020). We expect such mechanisms to efficiently clear out channels, either producing low-density channels or almost emptied holes with small clumps of matter remaining within. In both cases, we model such channels by a density-bounded component with low input ionization parameter (see Eq. (4)). Since the LyC emitters studied here are clearly dominated by very young stellar populations with ages  $\lesssim 3\text{--}4 \text{ Myr}$ , too young for significant feedback from supernovae (Izotov et al. 2018b), we suspect radiation pressure and related feedback effects to be the main driver for the creation of these channels or holes, although the explanation may be more complex (cf. Hogarth et al. 2020).

**Density-bounded galaxy (DBG).** Density-bounded galaxy with ionization-bounded clumps. The core of the galaxy is completely ionized by massive young stars, which produce a large number of ionizing photons exceeding what can be absorbed by the available nebular gas. This is represented by a high- $U$ , density-bounded component, from which photons can directly escape at the edge. We add a low-ionization component to represent denser, ionization-bounded filaments or small clumps, which can naturally emit the [O I] and [S II] emission lacking in one-zone models (see Sect. 3) and where the observed UV low-ionization lines are formed. Both components have the same chemical composition and the same ionizing spectrum, emitted by a central source.

For a first approach, we combined two arbitrary photoionization models with different ionization parameters: one with a high-ionization parameter  $U_h$  and one with a low-ionization parameter  $U_l$ . The high- and low-ionization parameters were

chosen “by hand” so that the nonleaking combined models accurately reproduce the upper part of the SDSS sequence on our diagnostic plots (see Sect. 4.2) and the tight SDSS sequence in the O32–O13 and O32–S23 plots (Figs. 5 and 9). We used  $U_h = 10^{-1}$  and  $U_l = 10^{-3.5}$ . We briefly comment on this choice and an optimizing procedure below (Sect. 5.1).

We computed both scenarios for an identical stellar population of 1 Myr, a metallicity  $12 + \log(\text{O}/\text{H}) = 8$ , and  $\text{N}/\text{O} = 0.1$ , as above. As our combined model is meant to represent a galaxy ionized by the same stellar population, we have to make sure that the photoionization models describing the two components also correspond to the same number of ionizing photons  $Q_h$  and  $Q_l$ , respectively. This is simply achieved by scaling the predicted line luminosities of the  $U_l$  model with the factor  $s = Q_h/Q_l$ . The two components are then combined with relative weights,  $\omega$  for the  $U_h$  component and  $(1-\omega)$  for the  $U_l$  component. Any ratio of emission lines A and B of the combined model then becomes

$$\frac{L_c^A}{L_c^B} = \frac{\omega L_h^A + (1-\omega)L_l^A s}{\omega L_h^B + (1-\omega)L_l^B s}, \text{ where } s = Q_h/Q_l. \quad (5)$$

For each scenario, we allowed the LyC escape fraction of the leaking component to vary from 0% to 80% in steps of 10%, provided by the 3MdB\_17 database. The total  $f_{\text{esc}}$  of the two-zone model, is computed consistently from the relative weight of zones  $\omega$ , by defining

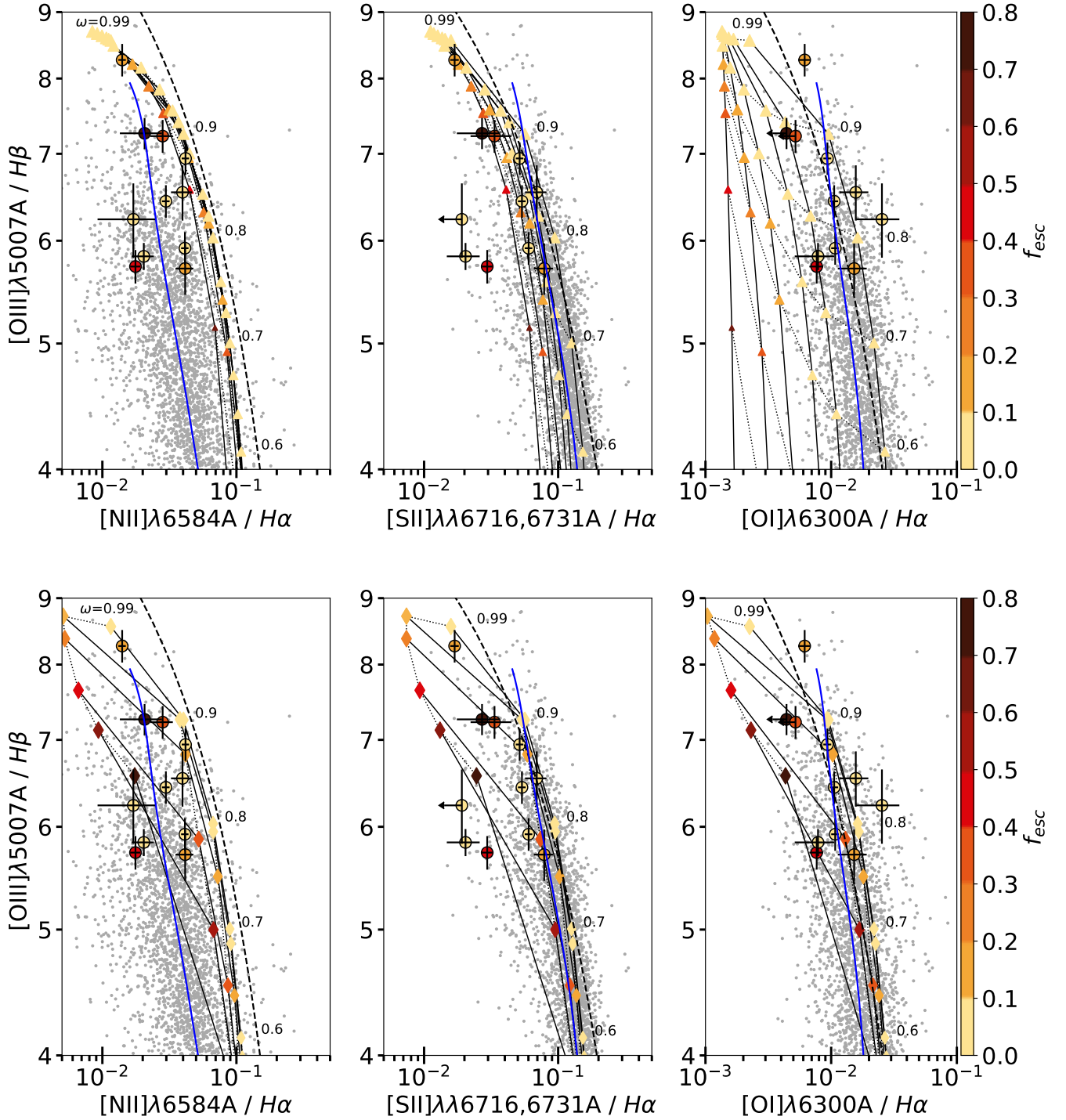
$$f_{\text{esc}}^c = \omega f_{\text{esc}}^h + (1-\omega)f_{\text{esc}}^l. \quad (6)$$

As only one component is leaking in each scenario, Eq. (6) simply becomes  $f_{\text{esc}}^c = (1-\omega)f_{\text{esc}}^l$  in the DBC scenario and  $f_{\text{esc}}^c = \omega f_{\text{esc}}^h$  in the DBG scenario. In the absence of leakage ( $f_{\text{esc}} = 0$ ), the two scenarios are trivially identical.

The resulting emission line ratios predicted from our two-zone models with varying relative weights and LyC escape fractions are shown in Figs. 8 and 9, allowing thus a comparison with the one-zone models (Figs. 4 and 5). Nonleaking models are represented by the yellow symbols with 0% escaping photons. In this case the two scenarios (DBC and DBG) are identical by construction, and their symbols (yellow triangles on the first row, and yellow diamonds on the second row) are thus superposed. As mentioned before, we chose the ionization parameter of each component such that the combined nonleaking models adequately cover the upper part of the SDSS sequence.

#### 4.2. [O I] and [S II] $\lambda\lambda 6716, 6731$ emission in normal galaxies explained

As we can see from Fig. 8, the nonleaking two-zone models follow a bended or sometimes straight “mixing line” from very high to moderate excitation (in [O III]  $\lambda 5007/\text{H}\beta$ ) and covering a wide range in the [N II]/H $\alpha$ , [S II]/H $\alpha$ , and [O I]/H $\alpha$  ratios. Despite our somewhat arbitrary choice of models describing the individual zones, we already see that observed line ratios of the bulk of the SDSS galaxy sample (which primarily hosts galaxies with very low or zero  $f_{\text{esc}}$ ) can be covered with the two-zone models combining two different ionization parameters. In particular, there seems to be no difficulty in explaining the relatively high observed intensities of [O I]  $\lambda 6300$ , and even the comparatively high values of [S II]  $\lambda\lambda 6716, 6731$ /[S III]  $\lambda 9068$  in normal galaxies, which cannot be reproduced by classical one-zone models (compare Figs. 5 and 9). The inclusion of a second ISM component with a lower ionization parameter therefore seems both essential and sufficient to solve the above-mentioned discrepancies of classical models when compared to normal galaxies.



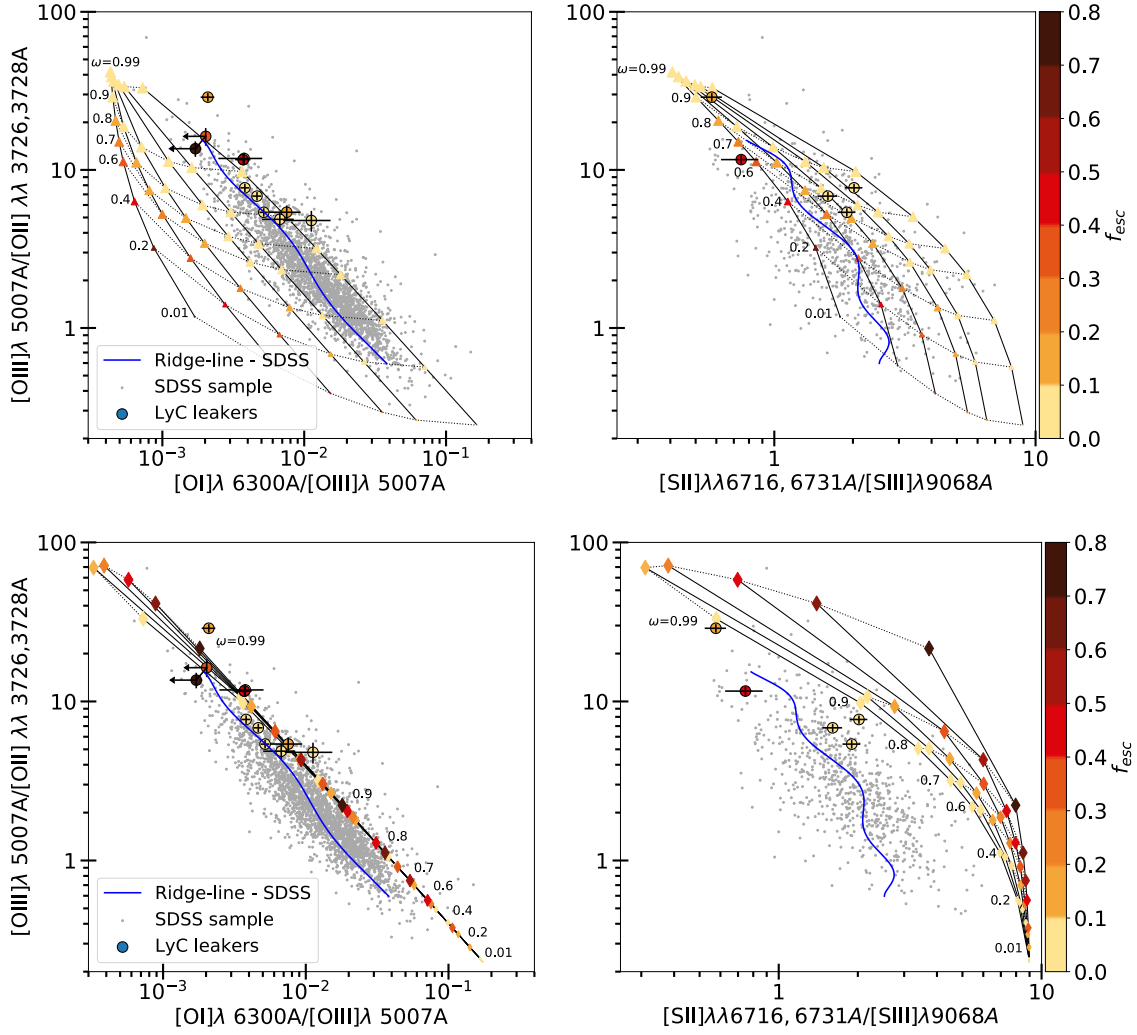
**Fig. 8.** Classical diagrams (same as Fig. 4) showing the predictions from the two-zone photoionization models. *Upper and lower panels:* represent DBC and DBG models, respectively.

#### 4.3. LyC emitters explained by two-zone models

In contrast to the one-zone models, the combined two-zone models are able to cover the entire range of emission line ratios of the LyC emitters, as shown by Figs. 8 and 9. Most importantly, the models can – at least qualitatively – explain the observed [O I] excess and the [S II] deficiency (see Figs. 1 and 2), which characterizes the leakers, and which cannot be understood by one-zone models (see Sects. 2.2 and 3.2). Furthermore, the two-zone models account for the escape of LyC photons, and are thus consistent with this important observational property. Below we

further explore the consistency of different line ratios and the extent to which the models also correctly reproduce the observed LyC escape fractions.

Let us now examine how the predictions from the two-zone models in classical diagrams vary with LyC escape fraction. In the O3H $\beta$ –N2H $\alpha$  diagram (Fig. 8 left), most of the models are superposed quite independently of their  $f_{\text{esc}}$  value. This is due to the fact that [N II] originates from the highly excited region in the inner part of the H II region, which remains unaffected when trimming the outskirts of the galaxy. Only the most



**Fig. 9.** O32–O13 and O32–S23 diagrams (same as Fig. 5) showing the predictions from the two-zone photoionization models. *Upper and lower panels:* DBC and DBG models, respectively.

density-bounded models from the DBG scenario, where the highly ionized component dominates almost all the galaxy ( $\omega = 99\%$ ) and is leaking, affect the  $[\text{N II}]/\text{H}\alpha$  ratio when  $f_{\text{esc}}$  exceeds 10%. In contrast, the  $[\text{S II}]$  and  $[\text{O I}]$  line ratios (shown in the middle and right panels) are very sensitive to LyC leakage. We note that these emissions primarily arise from the low-ionization component, that is, the one leaking in the DBC scenario (upper panel). This explains why these line ratios are barely affected in the DBG scenario (lower panel), even for high escape fractions. On the contrary, leakage in the low-ionization component (DBC scenario) rapidly decreases  $[\text{S II}]$  and  $[\text{O I}]$ : only low escape fractions ( $f_{\text{esc}} \lesssim 10\%$ ) are compatible with the observed  $[\text{O I}]$  emission of the leakers in this scenario.

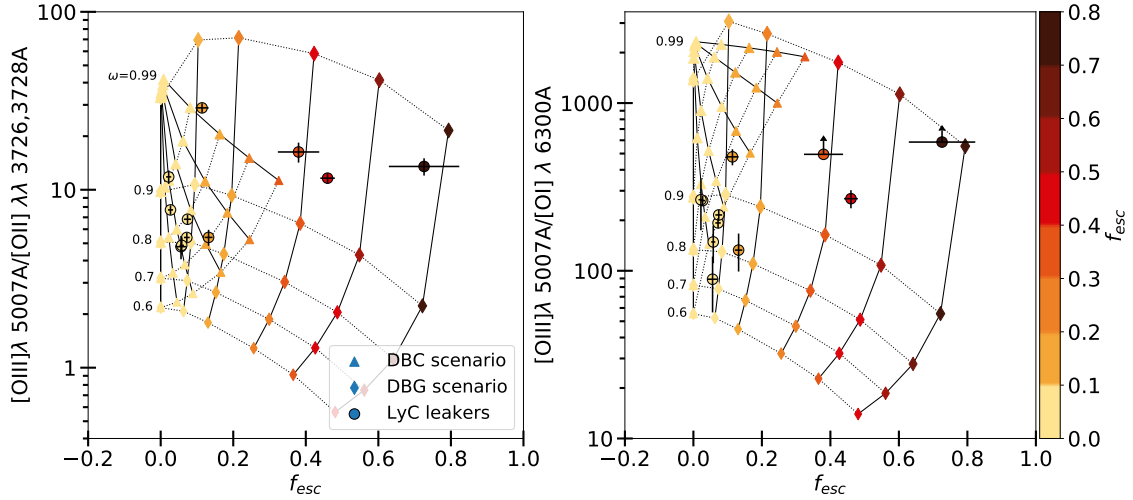
The same trend is visible on the O32–O13 plot (Fig. 9). The SDSS sequence seems bounded by  $f_{\text{esc}} \sim 0\text{--}10\%$  for the DBC scenario (first row). Such values are sufficient to explain the observed escape fraction of some of the leakers, represented by yellow circles, but it fails to reproduce higher leakage of ionizing photons, represented by the orange to brown circles. Meanwhile, the DBG scenario can produce escape fractions up to 80% without strongly decreasing the emission of  $[\text{O I}]$  and  $[\text{S II}]$ . On the other hand, the O32–S23 plan seems to be better covered by models from the DBC scenario, while the DBG scenario tends to overestimate the  $[\text{S II}]/[\text{S III}]$  ratio. For the five

leaking galaxies with  $[\text{S III}]$  measurements, the observed escape fractions seem consistent with predictions from the DBC scenario, even for the most leaky galaxy J1154+2443 (red circle,  $f_{\text{esc}} = 46\%$ ). Indeed, we note that this galaxy stands on the left-hand side of the median line, as expected for the highest escape fractions produced by the DBC scenario; it falls between two model points with escape fractions of 30–40%, which is roughly consistent with the observed value. We also note that models with a weight  $\omega \lesssim 0.7$  fail to reproduce the leakers studied here. This is easily understood if we keep in mind that these galaxies show high excitation, as indicated by their high O32 ratio: decreasing the contribution  $\omega$  of the  $U_{\text{high}}$  region below a certain threshold produces two-zone models that are dominated by the low ionization ( $U_{\text{low}}$ ) component, and hence are unable to reproduce the observed line ratios of the leakers.

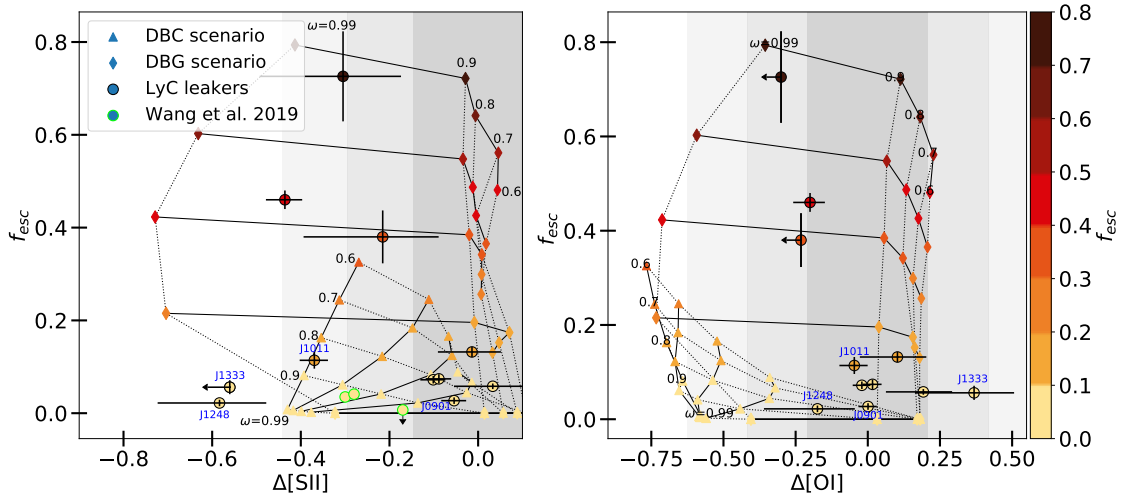
#### 4.4. Links with $f_{\text{esc}}$

In Fig. 10 we now revisit the predicted O32 and O31 line ratios in models with LyC escape presented earlier for one-zone models (see Fig. 6). As already mentioned and clearly shown in this figure, the two-zone models solve the  $[\text{O I}]$  problem of simple models, and they cover the entire range of observed line ratios of the LyC emitters. A comparison of the two panels also shows





**Fig. 10.** Same as Fig. 6 showing the predictions for the two-zone photoionization models.



**Fig. 11.** LyC escape fraction as a function of the deficiency  $\Delta[\text{S II}]$  in  $[\text{S II}]$  (left) and  $\Delta[\text{OI}]$  (right). The gray-shaded areas represent dispersion of the SDSS comparison sample at 1, 2, and 3 $\sigma$ . The three strongest leakers (J1256+4509, J1154+2443 and J1243+4646, in increasing order of  $f_{\text{esc}}$ ) are represented by orange to brown circles in the upper part of the plots. The other galaxies that are discussed in Sect. 4.4 are labeled in blue.

that, at least approximately and without adjustment of the components of the two-zone models, the models are capable of consistently reproducing the two oxygen line ratios and the observed escape fraction. Figure 10 also shows that LyC emitters with very high  $f_{\text{esc}} \gtrsim 20\%$  can only be reproduced by our DBG scenario (where the high  $U$  component is leaking), whereas both scenarios are possible for the galaxies with lower escape fractions.

Figure 11 (left) illustrates the evolution of  $f_{\text{esc}}$  as a function of  $\Delta[\text{S II}]$ . This plot suggests that a sufficiently strong  $[\text{S II}]$  deficiency (e.g.,  $\Delta[\text{S II}] \lesssim -0.2$ ) effectively picks out LyC leakers, although not all of them. Indeed, several LyC emitters are also located within the normal range of the comparison sample (presumably dominated by nonleakers), as also pointed out by Wang et al. (2019). However, using  $\Delta[\text{S II}]$  to determine the escape fraction is not obvious. Compared to the dispersion of our comparison sample represented by the gray-shaded areas, 6 out of 11 leakers show a deficit in  $[\text{S II}]$  larger than  $1\sigma$  (5 out of 11 with deficit larger than  $2\sigma$ ). If we include the 2 leaking galaxies from Wang et al. (2019), this amounts to 8 out of 13 galaxies with a clear  $[\text{S II}]$  deficit at  $1\sigma$  (6 out of 13 at  $2\sigma$ ). The three strongest leakers (J1256+4509, J1154+2443 and J1243+4646),

which have  $f_{\text{esc}} \gtrsim 30\%$  all exhibit a large deficiency in  $[\text{S II}]$  of at least 0.2 dex from the ridge line ( $<1\sigma$ ). J1011+1947, with an intermediate escape fraction of 11.4% also shows a large  $[\text{S II}]$  deficit. However, the most  $[\text{S II}]$ -deficient galaxy in our sample (J1248+4259) and J1333+6246, which is undetected in  $[\text{S II}]$ , have escape fractions of only a few percent. This could mean that the observed  $f_{\text{esc}}$  for these galaxies is underestimated, for example due to the fact that LyC radiation escapes only along other, unobserved lines of sight. Interestingly, in the analysis of the UV absorption lines of more than 30 low- $z$  galaxies from Gazagnes et al. (2020), J1248+4259 stands out as an outlier with an exceptionally strong Ly $\alpha$  emission (high equivalent width) despite showing a large covering fraction in the H I absorption lines. Demonstrating whether or not any of these sources indeed emit more ionizing photons in other directions seems difficult. In any case, our analysis confirms that the method proposed by Wang et al. (2019) could be a promising way to identify new leaking galaxies, although the determination of  $f_{\text{esc}}$  from the  $[\text{S II}]$  deficit remains challenging.

From Fig. 11 (left) we notice once more that only the DBG scenario is able to reproduce the more extreme leakage above

20%. The DBC scenario is able to cover the space where most of the galaxies with small  $f_{\text{esc}}$  reside, including the [S II]-deficient galaxies selected by Wang et al. (2019; green color along the contours of circles). However, we note that the two galaxies with the strongest [S II] emission ( $\Delta[\text{S II}]$  close to zero) stand outside of the contours of the DBC scenario. This indicates that leakage powered exclusively by channels of low ionization is unable to simultaneously produce strong [S II] emission and escaping LyC photons. This comes from the fact that in this scenario leakage is allowed exclusively by trimming the low ionization component, which always results in decreasing [S II] emission. This could either suggest that density-bounded LyC could still be the preferred mechanism even for rather small ( $\lesssim 10\%$ ) escape fractions, or that a more complex geometry needs to be considered to allow leakage without systematically decreasing [S II].

In Fig. 11 (right) we show  $\Delta[\text{O I}]$  calculated from the  $\text{O3H}\beta$ – $\text{O1H}\alpha$  plots and defined earlier (see Fig. 4, Sect. 2.2). Positive  $\Delta[\text{O I}]$  corresponds to galaxies showing an excess in [O I]/H $\alpha$  compared to the SDSS median line, while a negative  $\Delta[\text{O I}]$  marks an [O I] deficiency. Interestingly, we notice a visible trend of  $f_{\text{esc}}$  relative to  $\Delta[\text{O I}]$  in the observations: while the most [O I]-deficient galaxies and the two undetected galaxies ( $|\Delta[\text{O I}]| < 1\sigma$ ) exhibit the largest escape fractions, the observed  $f_{\text{esc}}$  decreases for increasing  $\Delta[\text{O I}]$ . This could indicate that [O I] emission, which traces the photo-dissociated interface between ionized and atomic gas or cold dense clumps in the H II region might be linked to the disappearance of the remaining Lyman continuum, leading to lower escape fractions. However, we notice that two galaxies (J1248+4259 and J0901+2119) that are [O I]-deficient exhibit very small escape fractions ( $f_{\text{esc}} \sim 2\%$ ). This indicates that other effects are likely to cause the drop of Lyman continuum even in [O I]-deficient environment. However, the overall trend shown by the observations in Fig. 11 suggests that selecting [O I]-deficient galaxies could also be an effective way to target galaxies with high leakage and might correlate better with  $f_{\text{esc}}$  than [S II] deficiency. On the other hand, the [S II] lines are generally stronger than [O I], making the former easier to measure.

In Fig. 11 (right) the contours from the DBC scenario include only the two galaxies with the lowest escape fractions. For the same reasons as for [S II] emission, leakage through low-ionization channels tends to be associated with low [O I] emission, that is,  $\Delta[\text{O I}] \lesssim 0$ . The galaxies that are not [O I] deficient are not well reproduced by the DBC scenario, especially the galaxy standing on the extreme right-hand side whose [O I] emission exceeds that of all the models considered here. This might suggest that the physics of [O I] emission is not well taken into account in our two-component models, and might require a boost from an additional component (see discussion in Sect. 5). However, we note that apart from the systematic offset toward lower [O I] emission, the DBC scenario qualitatively reproduces the observed trend of increasing  $f_{\text{esc}}$  with decreasing  $\Delta[\text{O I}]$ . The DBG scenario accurately reproduces all the  $\Delta[\text{O I}]$  and  $f_{\text{esc}}$  of the leakers (except for the strongest [O I] emitter) but  $\Delta[\text{O I}]$  is independent from the escape fraction in this scenario; the observed variations in  $\Delta[\text{O I}]$  only come from the decrease of [O III] which vertically moves the models closer to the ridge line but without affecting [O I] emission.

We conclude that  $\Delta[\text{S II}]$  and  $\Delta[\text{O I}]$  could be interesting proxies to estimate the escape fraction but only in the case of ionization-bounded leakage (DBC scenario). In the case of density-bounded leakage (DBG scenario), likely producing the highest escape fractions, [O I] and [S II] seem uncorrelated with  $f_{\text{esc}}$ , and other diagnostics sensitive to density boundaries are to

be preferred (see, e.g., Figs. 5 and 9). It seems that [O III]/[O II],  $\Delta[\text{S II}]$ , and [O I] excess are good hints for LyC leakage, but there still seems to be other independent dimensions, which probably incorporate several physical parameters such as the nature of the ionizing cluster, its evolutionary stage, the gas content of each galaxy, its star formation rate, and metallicity, all of which could produce other combinations for a LyC leaking cluster (e.g., Wang et al. 2019).

## 5. Discussion

We now briefly discuss further improvements, potential caveats, implications, and future prospects resulting from the present work.

### 5.1. Can the LyC escape fraction be detected and measured from the optical emission lines?

In this study, we used the observed LyC escape fractions and optical emission lines to build simple but consistent two-zone photoionization models of the LyC emitters, and to discriminate between different leakage scenarios. An obvious further step would be to use these two-component models to infer the expected escape fraction from the emission line observations only. Such indirect methods are badly needed in order to optimally exploit upcoming JWST observations of rest-frame optical emission lines for galaxies during the epoch of reionization.

A first automated routine was implemented in an effort to extract best-fit values for the parameters of the two-component models. Although very promising, this method requires a robust implementation and a careful choice of tracers, which we postpone to future work. We present here the first trends obtained from this study and discuss possible caveats of this method.

We ran a  $\chi^2$ -minimization routine on a subgrid of our combined models, which includes the ones presented in Sect. 4. We defined a four-dimensional grid with the following free parameters:  $U_{\text{high}}$  and  $U_{\text{low}}$  (varied from  $10^{-4}$  to  $10^{-1}$ ), the weight of the component of high ionization  $\omega$ , and the escape fraction of the leaking component  $f_{\text{esc}}^{\text{holes}}$  (both varied between 0 and 1). As observational constraints we used six line ratios: the four ratios used in the classical diagrams (Fig. 8), and two ratios from the O32–O13 diagrams (Fig. 9). We did not include upper limits. As our prescription for  $f_{\text{esc}}^{\text{combined}}$  is fully determined by the parameters of each model (see Eq. (6)), we could also determine the corresponding value of the total LyC escape fraction for the  $\chi^2$  minimum. We also examined the weight of the ionization-bounded component  $\omega$  in the DBC scenario (respectively  $(1 - \omega)$  in the DBG scenario) to compare it to the UV covering fractions deduced from the Lyman series absorption lines (Gazagnes et al. 2018; Chisholm et al. 2018; Schaerer et al. 2018). Hereafter, we summarize some interesting trends that merit further examination:

- The models minimizing  $\chi^2$  always favor two distinct values for the ionization parameters  $U_{\text{high}}$  and  $U_{\text{low}}$ . Inspection of their marginalized probability distribution functions shows that typically  $U_{\text{high}} \sim 10^{-1} - 10^{-2}$ , and  $U_{\text{low}} \sim 10^{-3} - 10^{-4}$ . We interpret this as confirmation of our conclusion in Sect. 4 that two-component models do a better job at reproducing the observed line ratios than single-component models.

- Regardless of the considered scenario, the weight of the high-ionization component is always much higher than that of the low-ionization one ( $\omega \sim 0.8$ ). This means that no matter which component is leaking, the high-ionization component always dominates the total emission, in the same proportion

of about 80% (respectively 20% for the low-ionization component). This suggests that the evolution from one scenario to the other could result from a continuous evolution (cf. Kakiichi & Gronke 2019) leaving the global proportions of the structure of the galaxy unchanged. However, we note that the two scenarios produce very different covering fractions. We note that for galaxies with a low escape fraction ( $f_{\text{esc}} \lesssim 10\%$ ), the  $\omega$  obtained from the best-fitting model has a value comparable to the covering fractions,  $C_f$ , obtained from analysis of absorption lines (Gazagnes et al. 2018, 2020; Chisholm et al. 2018). This is consistent with what we expect from the DBC scenario, where the covering fraction should be close to  $\omega$  but cannot be reconciled with the DBG scenario, where  $C_f \sim (1 - \omega)$ , which is a better fit for some of these galaxies. On the other hand, for strong leakers ( $f_{\text{esc}} \gtrsim 10\%$ ), we find  $C_f < \omega$ . This strengthens the conclusion in Sect. 4.4 that the DBC scenario is unable to reproduce strong leakage, as it predicts too high a covering fraction, and therefore overly low escape fractions. However, we note that the DBC scenario sometimes better reproduces the six line ratios we considered, even for strong leakers.

– Finally, the escape fraction obtained for the best-fit models allows us to distinguish the four galaxies with the strongest leakage in our sample. The best-fit models predict  $f_{\text{esc}} \approx 0$  for all the weak leakers but one, and nonzero escape fractions ( $f_{\text{esc}} \sim 0.6\text{--}38\%$ ) for the galaxies with high escape fractions. We note that the highest predicted escape fraction corresponds to J1243+4646, the strongest leaker in our sample, although the predicted value of  $f_{\text{esc}}$  is a factor of approximately two below the observed value.

Briefly, this exercise supports our conclusion in Sect. 4.4 that the geometry of weak leakers resembles that of the DBC scenario, and strong leakers more the DBG scenario. However, we find that robustly distinguishing between the two scenarios based only on emission lines is not easy. Furthermore, accurate predictions of  $f_{\text{esc}}$  from the optical emission line ratios do not appear to be straightforward. This will probably require a careful choice of the relevant line fluxes or ratios to be used in such an analysis, a proper treatment of upper limits and uncertainties, a robust statistical method to overcome the caveats of a simple  $\chi^2$  minimization, the exploration of SEDs allowing also for different ages, and other improvements, which are beyond the scope of the present publication.

## 5.2. Low ionization line diagnostics: progress and caveats

In Sect. 3, we discussed the inadequacy of simple models to simultaneously reproduce the high and low ionization lines, as revealed for example in the O32–O13 and O32–S23 diagnostics (Fig. 5). Specifically, we emphasized the offset of BOND models in O32–S23, which systematically underestimate the [S II]/[S III] ratio. This offset of standard photoionization models has been shown in previous work (e.g., Pérez-Montero & Vílchez 2009; Kehrig et al. 2006), in which such models fail to explain the hardening in ionizing radiation shown by most low-mass, low-metallicity H II galaxies in the O32–S23 plane. Problems with the sulfur line strengths have often been reported in the past (see, e.g., Garnett 1989; Kewley et al. 2019; Mingozi et al. 2020, and references therein). The shortcomings of several Cloudy and Mappings models to reproduce the [S II]/[S III] ratio have for example recently been shown by Mingozi et al. (2020). Considering that this only seems to affect sulfur lines, these latter authors suggest that this is due to limitations in stellar atmosphere modeling and/or the atomic data for sulfur. Based on the observation that the discrepancy holds for single H II regions

(from Kreckel et al. 2019), Mingozi et al. (2020) exclude the role of diffuse ionized gas as a cause of this issue. However, in our study, we see in Fig. 9 that adding a low-ionization component, which provides additional emission of [S II] without producing [S III], helps to solve this discrepancy. We conclude that for both our reference sample of star-forming galaxies and the LyC-emitting galaxies, the addition of such a low-ionization component is sufficient to reproduce the observed [S II]/[S III] ratios. Furthermore, this component also contributes to low ionization emission of oxygen, simultaneously reproducing the observed line strength of [O I]  $\lambda 6300$  and lines tracing the other ionization stages of oxygen. However, the exact nature, spatial distribution, and other properties of such a component remain to be understood.

While this success of our two-zone models is notable and certainly calls for further studies and applications, it must be noted that other physical processes can also play a role for the low-ionization lines. For example, it is well known that shocks in the ISM or the presence of sources with hard ionizing spectra (far-UV and X-ray emission) lead to enhanced emission of highly ionized species (e.g., [Ne V], [Fe V], [O IV], He II, and others) and also of low-ionization lines (e.g., Schmitt 1998; Stasińska et al. 2015; Plat et al. 2019). The effect of a contribution from shock models with different magnetic field strengths (tuned to reproduce a typical He II/H $\beta$  ratio) on the O32–O13 diagram is shown in Stasińska et al. (2015). Clearly, some model combinations of shock with both density- and ionization-bounded H II regions can be found that reproduce the three ionization stages of oxygen. How these models would affect the sulfur lines, and whether or not they could also be consistently reproduced remains to be seen. Given different degeneracies and uncertainties it seems currently difficult to exclude a contribution from shocks (see Stasińska et al. 2015). Similarly, a weak AGN component (contribution  $\sim 3\%$  of the hydrogen ionizing photons) cannot be excluded based on the observed emission line diagrams for the galaxies of our reference sample (Stasińska et al. 2015). Again, the sulfur line ratios used here may provide an additional handle on these questions.

Finally, the low ionization lines are also affected by diffuse ionized gas (DIG) in galaxies, as recently discussed by Mingozi et al. (2020), Shapley et al. (2019), Sanders et al. (2020), and others. In the LyC emitters studied here, the contribution of the DIG would presumably be low, if we adopt the observed scaling (e.g., Oey et al. 2007; Shapley et al. 2019), since the LyC emitters are compact by selection, and have a very high SFR surface density (Izotov et al. 2018b).

Regardless of the exact origin of emission of the low-ionization lines (such as [O I]  $\lambda 6300$  and [S II]  $\lambda\lambda 6716, 6731$ ) in LyC emitting galaxies, the UV observations of these galaxies clearly reveal the presence of both low-H I-column-density ( $N_{\text{HI}}$ ) lines of sight towards the observer allowing the escape of LyC radiation and lines of sight with high  $N_{\text{HI}}$ , where other neutral and low-ionization metal absorption lines are also formed (Gazagnes et al. 2018, 2020; Chisholm et al. 2018). Therefore, their ISM contains at least two “zones” or regions with different conditions, even if we do not know the exact spatial scales on which this happens. Furthermore, it is very plausible that properties that influence the ionization parameter (e.g., density, clumpiness, filling factor) differ between these regions. This justifies our two-zone approach using two different ionization parameters,  $U_{\text{high}}$  and  $U_{\text{low}}$ , for the LyC leakers. Whether or not and to what extent such an approach is also motivated for “normal” galaxies (most of which are presumably not LyC emitters)



remains to be seen. In any case, given the ubiquitous complexity of the ISM in resolved star-forming galaxies it is probably not surprising that simple 1D photoionization models cannot reproduce all the observables.

### 5.3. Future improvements

Some of the methodology used and some of the questions addressed here certainly call for other applications and for future improvements. Indeed, although photoionization models are frequently used to predict and analyze emission lines in star-forming galaxies, they are generally 1D spherical or plane parallel models computed with Cloudy or similar codes, which then allow for exploration in a wide range of parameter space in terms of ionization parameters, radiation field, abundances and so on. Such models, which typically assume a single source of ionizing radiation and a very simple geometry, in contrast to observations of integrated galaxy spectra, are obvious oversimplifications of the true complexity and diversity of the radiation field, ISM, and geometry.

Relatively few more highly complex models have so far been used for comparisons with observations. Stasińska et al. (2015) for example explored the combination of stellar photoionization and other emission sources (AGN, shocks, and others) with variable but arbitrary contributions, and confronted them with a large sample of star-forming galaxies from the SDSS. Cormier et al. (2012, 2019) used one- and two-component models to consistently analyze emission lines originating in the ionized and neutral ISM, as well as in adjacent PDRs. These have been compared to detailed multi-wavelength observations of 39 nearby dwarf galaxies. A very complex tailored model including up to five components, allowing also for ionization- and density-bounded parts, has been constructed for the well-known metal-poor galaxy I Zw 18 by Leboutteiller et al. (2017). Clearly other models can and need to be envisaged, although sufficient independent observational constraints are necessary to avoid degeneracies in a multi-dimensional parameter space.

Another approach is to use high-resolution hydrodynamic simulations of galaxies, ideally including radiation transfer and a proper treatment of the ISM, to predict the resulting emission lines and compare them to observations. Such models, with fairly different degrees of sophistication have for example been presented by Shimizu et al. (2016), Hirschmann et al. (2017), Katz et al. (2019), Wilkins et al. (2020), and by Pellegrini et al. (2020) for smaller scales. These simulations are able to capture several observational trends and allow useful predictions, in particular for high-redshift galaxies. However, direct and detailed comparisons to observations are not trivial to undertake or to interpret. For example, Katz et al. (2020) predicts optical and IR emission lines of LyC leakers and nonleakers, which show good agreement for the [O III]  $\lambda\lambda 4959, 5007$ , [O II]  $\lambda\lambda 3726, 3728$ , and H $\beta$  emission lines when compared to observations. In contrast, their predictions for [S II]  $\lambda\lambda 6716, 6731/H\alpha$  and [O III]/H $\beta$  are quite different from those observed in low- $z$  LyC emitters and in  $z \sim 2-3$  galaxies. Understanding the origin and importance of such a discrepancy is nevertheless not straightforward. Different and complementary modeling approaches should help to develop more robust diagnostics and should therefore provide better insight into the properties of the ISM, radiation field, geometry, and other properties of star-forming galaxies and LyC emitters.

In parallel, new observations, such as those from the HST Low- $z$  LyC survey (PI A. Jaskot), will provide for the first time

a statistical sample of more than 60 galaxies at  $z \sim 0.2-0.4$  with direct LyC observations. This sample of confirmed LyC emitters and also nonLyC emitters should represent a useful base to test several of the trends observed here, and an ideal sample for detailed comparisons with emission line models from different approaches. The first comparison with simple two-zone models and a small sample presented here is very encouraging, but incomplete. Nevertheless, there is hope that optical emission lines can to some extent be used to identify LyC emitters, and will ideally allow determination (even approximate) of the escape fraction of ionizing photons from these galaxies. Clearly, future improvements on these lines will be very welcome and of great interest, especially in view of the upcoming launch of the JWST, which will routinely observe rest-frame optical emission lines out to the highest redshifts.

## 6. Conclusion

We analyzed the main optical emission lines of low- $z$  LyC-emitting galaxies (leakers) discovered recently with the HST (Izotov et al. 2016a,b, 2018a,b; Wang et al. 2019) in the classical excitation diagrams involving the [O III]/H $\beta$ , [N II]/H $\alpha$ , [S II]/H $\alpha$ , and [O I]/H $\alpha$  line ratios, as well as other diagnostic diagrams involving three ionization stages of oxygen (using [O III], [O II], and [O I]) and two of sulfur (using [S II] and [S III] observations). In comparison with the SDSS sample of star-forming galaxies with similar metallicities ( $12 + \log(\text{O}/\text{H}) \sim 7.6-8.2$ ), we found several peculiarities of the LyC emitters, including an excess of the [O I]/[O III] ratio for a fraction of the leakers (as already pointed out by Plat et al. 2019), and an [O I] deficiency for those with the highest LyC escape fractions ( $f_{\text{esc}} \gtrsim 20\%$ ). We also confirm a deficit in [S II] found recently by Wang et al. (2019), and suggest that combined measurements of the [O I] and [S II] lines could provide a good indicator for strong LyC emission.

We then compared the observations to simple one-zone and two-zone photoionization models accounting for the escape of ionizing radiation in order to construct, for the first time, consistent models which reconcile their observed emission lines and the measured LyC escape fractions. The main results from this analysis can be summarized as follows:

- Classical one-zone photoionization models underpredict the low-excitation line emission of [O I] and [S II] arising from the outskirts of H II regions; this underprediction is strong for density-bounded models with LyC escape, and also exists for ionization-bounded models (see Sect. 3), although it is less strong.

- Two-zone models, combining regions with a high- and low-ionization parameter ( $U_{\text{high}}$ ,  $U_{\text{low}}$ ), where one of which is density-bounded, allowed us to resolve these discrepancies and to reproduce the observed line ratios of the LyC emitters. Furthermore, the models show LyC escape fractions comparable to those measured directly from the UV LyC observations.

- From the two-zone models we also found a possible dichotomy between galaxies with different LyC escape fractions. Indeed, the models indicate that galaxies with  $f_{\text{esc}} < 10\%$  can be explained by the presence of low-column-density, low-filling-factor regions or channels with a low-ionization parameter through which the ionizing photons escape. The observed [O I] emission can originate from these filaments/channels, in what we called the DBC scenario (Fig. 7). For much higher  $f_{\text{esc}}$ , our photoionization models show that LyC leakage must occur from the region with high-ionization parameter (DBG scenario) in order to explain the observed emission line ratios.

Our two-zone models are inspired by and compatible with the “picket-fence”-type model, which successfully reproduces the observed UV absorption lines of hydrogen and low-ionization metal lines of the LyC emitters, as demonstrated by the studies of [Gazagnes et al. \(2018, 2020\)](#). Our finding of two different scenarios, favored by LyC emitters with a low or high escape fraction, is also compatible with the recent analysis of [Gazagnes et al. \(2020\)](#) combining information from the Ly $\alpha$  line profiles and absorption lines. Our results are also consistent with the numerical simulations of [Kakiichi & Gronke \(2019\)](#), which suggest that the mechanism powering LyC escape may evolve in a continuous way from the DBC scenario, producing low escape fractions, to the DBG scenario, which appears as a necessary configuration to reach high levels of leakage.

Finally, we have discussed the possibility of improved analysis tools to predict escape fractions based on optical emission lines and using two-zone photoionization models (Sect. 5). We conclude that robust predictions would require multi-phase and multi-sector modeling, including some density-bounded components to successfully capture the complexity of the ISM structure that leads to photon leakage. We also suggest that low-excitation tracers from the outer part of H II regions such as [O I] and [S II] lines should be included in the list of relevant lines to study the morphology and dynamical state of leaking and nonleaking galaxies.

**Acknowledgements.** We acknowledge the use of photoionization models from Stephane de Barros at an earlier stage of this work. YII and NGG acknowledge support from the National Academy of Sciences of Ukraine by its priority project No.0120U100935 “Fundamental properties of the matter in the relativistic collisions of nuclei and in the early Universe”. JMV acknowledges support from the State Agency for Research of the Spanish MCIU through the “Center of Excellence Severo Ochoa” award to the IAA (SEV-2017-0709) and project AYA2016-79724-C4-4-P. RA acknowledges support from FONDECYT Regular Grant 1202007.

## References

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, **235**, 42
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, **93**, 5
- Bassett, R., Ryan-Weber, E. V., Cooke, J., et al. 2019, *MNRAS*, **483**, 5223
- Bian, F., Fan, X., McGreer, I., Cai, Z., & Jiang, L. 2017, *ApJ*, **837**, L12
- Chabrier, G. 2003, *ApJ*, **586**, L133
- Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018, *A&A*, **616**, A30
- Cormier, D., Leboutteiller, V., Madden, S. C., et al. 2012, *A&A*, **548**, A20
- Cormier, D., Abel, N. P., Hony, S., et al. 2019, *A&A*, **626**, A23
- de Barros, S., Vanzella, E., Amorín, R., et al. 2016, *A&A*, **585**, A51
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *Rev. Mex. Astron. Astrofis.*, **53**, 385
- Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, *ApJ*, **878**, 87
- Garnett, D. R. 1989, *ApJ*, **345**, 282
- Gazagnes, S., Chisholm, J., Schaerer, D., et al. 2018, *A&A*, **616**, A29
- Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A&A*, **639**, A85
- Guseva, N. G., Izotov, Y. I., Fricke, K. J., & Henkel, C. 2019, *A&A*, **624**, A21
- Guseva, N. G., Izotov, Y. I., Schaerer, D., et al. 2020, *MNRAS*, **497**, 4293
- Hirschmann, M., Charlot, S., Feltre, A., et al. 2017, *MNRAS*, **472**, 2468
- Hogarth, L., Amorín, R., Vílchez, J. M., et al. 2020, *MNRAS*, **494**, 3541
- Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, *A&A*, **448**, 955
- Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016a, *Nature*, **529**, 178
- Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, *MNRAS*, **461**, 3683
- Izotov, Y. I., Thuan, T. X., & Guseva, N. G. 2017, *MNRAS*, **471**, 548
- Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, **474**, 4514
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, **478**, 4851
- Izotov, Y. I., Thuan, T. X., & Guseva, N. G. 2019, *MNRAS*, **483**, 5491
- Jaskot, A. E., & Oey, M. S. 2013, *ApJ*, **766**, 91
- Kakiichi, K., & Gronke, M. 2019, *ApJ*, submitted [arXiv:1905.02480]
- Katz, H., Galligan, T. P., Kimm, T., et al. 2019, *MNRAS*, **487**, 5902
- Katz, H., Đurovčiková, D., Kimm, T., et al. 2020, *MNRAS*, **498**, 164
- Kehrig, C., Vílchez, J. M., Telles, E., Cuisinier, F., & Pérez-Montero, E. 2006, *A&A*, **457**, 477
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, **142**, 35
- Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001, *ApJS*, **132**, 37
- Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, *ARA&A*, **57**, 511
- Kreckel, K., Ho, I. T., Blanc, G. A., et al. 2019, *ApJ*, **887**, 80
- Leboutteiller, V., Péquignot, D., Cormier, D., et al. 2017, *A&A*, **602**, A45
- Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, *ApJ*, **823**, 64
- McLeod, A. F., Dale, J. E., Ginsburg, A., et al. 2015, *MNRAS*, **450**, 1057
- Mingozzi, M., Belfiore, F., Cresci, G., et al. 2020, *A&A*, **636**, A42
- Mollá, M., García-Vargas, M. L., & Bressan, A. 2009, *MNRAS*, **398**, 451
- Morisset, C., Delgado-Inglada, G., & Flores-Fajardo, N. 2015, *Rev. Mex. Astron. Astrofis.*, **51**, 103
- Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. 2018, *MNRAS*, **478**, 791
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, **442**, 900
- Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, **889**, 161
- Oey, M. S., Meurer, G. R., Yelda, S., et al. 2007, *ApJ*, **661**, 801
- Pellegrini, E. W., Rahner, D., Reissl, S., et al. 2020, *MNRAS*, **496**, 339
- Pérez-Montero, E., & Vílchez, J. M. 2009, *MNRAS*, **400**, 1721
- Plat, A., Charlot, S., Bruzual, G., et al. 2019, *MNRAS*, **490**, 978
- Sanders, R. L., Shapley, A. E., Zhang, K., & Yan, R. 2017, *ApJ*, **850**, 136
- Sanders, R. L., Jones, T., Shapley, A. E., et al. 2020, *ApJ*, **888**, L11
- Schaerer, D., Izotov, Y. I., Verhamme, A., et al. 2016, *A&A*, **591**, L8
- Schaerer, D., Izotov, Y. I., Nakajima, K., et al. 2018, *A&A*, **616**, L14
- Schmitt, H. R. 1998, *ApJ*, **506**, 647
- Shapley, A. E., Sanders, R. L., Shao, P., et al. 2019, *ApJ*, **881**, L35
- Shimizu, I., Inoue, A. K., Okamoto, T., & Yoshida, N. 2016, *MNRAS*, **461**, 3563
- Stasińska, G., & Leitherer, C. 1996, *ApJS*, **107**, 661
- Stasińska, G., & Schaerer, D. 1999, *A&A*, **351**, 72
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, *MNRAS*, **371**, 972
- Stasińska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, *A&A*, **576**, A83
- Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, *ApJ*, **869**, 123
- Vale Asari, N., Stasińska, G., Morisset, C., & Cid Fernandes, R. 2016, *MNRAS*, **460**, 1739
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, **825**, 41
- Vanzella, E., Nonino, M., Cupani, G., et al. 2018, *MNRAS*, **476**, L15
- Vanzella, E., Caminha, G. B., Calura, F., et al. 2020, *MNRAS*, **491**, 1093
- Veilleux, S., & Osterbrock, D. E. 1987, *ApJS*, **63**, 295
- Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, *A&A*, **578**, A7
- Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, **597**, A13
- Wang, B., Heckman, T. M., Leitherer, C., et al. 2019, *ApJ*, **885**, 57
- Westmoquette, M. S., Smith, L. J., Gallagher, J. S., & Walter, F. 2013, *MNRAS*, **428**, 1743
- Wilkins, S. M., Lovell, C. C., Fairhurst, C., et al. 2020, *MNRAS*, **493**, 6079